Chapter 4: Changes in Services for Preschoolers With Disabilities

Special education services for preschool-aged children may include a wide array of instructional and related services. In PEELS, teachers were asked to identify all services provided to the participating children through the school system. A list of 23 services was included as well as a place to enter other services[21]. Teachers indicated that nearly 90 percent (89%, *S.E.* = 1.5) of the children received speech or language therapy in the 2003-04 school year, and 86 percent (*S.E.* = 1.6) did so in 2004-05, making it the most common service both years. Between one-fifth and one-third of children received occupational therapy (32%, *S.E.* = 1.6 in 2003-04 and 35%, *S.E.* = 1.5 in 2004-05) and learning strategies/study skills assistance by a special educator (30%, *S.E.* = 2.0 in 2003-04 and 20%, *S.E.* = 1.0 in 2004-05).

In this chapter, the augmented sample was used for all analyses. Also, all analyses in this chapter were longitudinal. T-tests for dependent samples were used to assess statistically significant changes in services over 2 years within subgroups. No analyses were conducted to compare differences across the subgroups. To control the family-wise error rate and avoid making false positive claims, the Benjamini-Hochberg procedure (Benjamini and Hochberg 1995) was used for all multiple testing situations.

On average, preschoolers with disabilities in PEELS were nearly 3 years old when they started receiving special education or therapy services from a professional (Markowitz et al. 2006). As children get older, one might expect their services to change, either to reflect the emerging demands of an elementary-school setting or to reflect children's evolving educational and related service needs. From 2003-04 to 2004-05, teachers reported a decrease in the percentage of children receiving many services. It should be noted that although the outcome variable is categorical, with large sample sizes, changes in proportion can be assumed to have a normal *t* distribution, so a *t* test is appropriate.

There was a statistically significant reduction in the number of children receiving nine services (see table 19). For example, service coordination/case management decreased from 25 percent (*S.E.* = 2.4) in 2003-04 to 9 percent (*S.E.* = 1.0) in 2004-05; training, counseling, or other supports/services for the children's family decreased from 16 percent (*S.E.* = 1.5) in 2003-04 to 5 percent (*S.E.* = 0.6) in 2004-05; and learning strategies/study skills assistance decreased from 30 percent (*S.E.* = 2.0) in 2003-04 to 20 percent (*S.E.* = 1.0) in 2004-05. The only service showing a significant increase from 2003-04 to 2004-05 was help from a one-to-one paraeducator or assistant, which increased from 10 percent (*S.E.* = 0.8) to 13 percent (*S.E.* = 1.1).

In 2003-04, young children with disabilities, on average, received 3.5 different special education and related services (*S.E.* = 0.1); in 2004-05, that figure was 2.8 (*S.E.* = 0.1) (see table 20). From 2003-04 to 2004-05, the mean number of services received by children decreased significantly overall and for children in each cohort, for males and females, and for all disability categories except other health impairment.

[21] This list did not include indirect services, such as teacher consultation.

Table 19. The percentage of young children who received preschool special education services that received specific services through their school system, by school year: School years 2003-04 and 2004-05

	2003-04	2004-05	*t* value	*p* value
Adaptive physical education	9.6	11.0	1.261	0.212
Assistive technology services/devices	10.1	8.5	-0.91	0.366
Audiology*	9.7	4.2!	-2.76	0.008
Augmentative or alternative communication system*	10.0	6.5	-3.198	0.002
Behavior management program	14.4	11.9	-1.948	0.056
Learning strategies/study skills assistance*	29.5	20.4	-4.701	0
Occupational therapy*	31.9	35.4	2.215	0.03
One-to-one paraeducator/assistant*	9.8	13.0	2.548	0.013
Physical therapy	17.6	17.2	-0.377	0.707
Service coordination/case management*	25.4	8.6	-7.872	0
Social work services*	8.7	4.9	-3.377	0.001
Special transportation because of disability*	19.0	13.2	-4.384	0
Specialized computer software or hardware	6.4	5.0	-1.195	0.237
Speech or language therapy	88.6	86.4	-1.146	0.256
Training, counseling, or other supports/services for family*	16.4	4.5	-7.013	0
Tutoring/remediation by a special education teacher*	16.8	10.7	-4.013	0
Other services	17.0	13.6	-2.002	0.05

! Interpret data with caution.
* t, $p < .05$.
NOTES: Other services include health services; instruction in American Sign Language, Manual English, Cued Speech, or Braille; mental health services; reader or interpreter; vision services; and other services specified by the respondent. Denominators do not include children who were declassified from special education, so percentages include only children with an IEP.
SOURCE: U.S. Department of Education, National Center for Special Education Research, Pre-Elementary Education Longitudinal Study (PEELS), "Elementary School Teacher Questionnaire," "Kindergarten Teacher Questionnaire," and "Early Childhood Teacher Questionnaire."

Time Spent in Different Educational Settings

Some young children with disabilities receive relatively low-intensity services, for example, an hour per week of speech therapy. Others receive services for an entire school day. They may include all-day public or private preschool, kindergarten, or elementary school programs.

Teachers in PEELS were asked to indicate the amount of time per week the children spent in a variety of settings, including a regular education classroom, special education setting (e.g., resource room, self-contained class, or separate school for children with disabilities), therapy setting, non-special education setting outside the classroom for remedial or special assistance (e.g., resource room for Title I or English-as-a-second-language instruction), and home instruction. From 2003-04 to 2004-05, the mean hours per week spent in a regular education classroom increased significantly, from 8.2 hours (*S.E.* = 0.5) to 15.0 hours (*S.E.* = 0.5) (t = 15.535, $p < .001$). Time in special education settings decreased significantly, from 8.0 hours a week in 2003-04 (*S.E.* = 0.5) to 6.2 hours a week in 2004-05 (*S.E.* = 0.4) (t = -4.315, $p < .001$). The amount of time children spent in a therapy setting, non-special education

Table 20. Mean number of services provided to young children who received preschool special education services, by age cohort, gender, and disability: School years 2003-04 and 2004-05

	2003-04	2004-05	*t* value	*p* value
Total	3.5	2.8	-6.03	0
Age cohort				
Cohort A*	3.6	2.7	-6.74	0
Cohort B*	3.6	2.7	-5.65	0
Cohort C*	3.3	2.8	-2.72	0.008
Gender				
Male*	3.5	2.8	-6.28	0
Female*	3.5	2.7	-3.77	0
Primary disability at Wave 1				
Autism*	5.5	4.1	-2.93	0.005
Developmental delay*	4.2	3.2	-5.48	0
Emotional disturbance*	4.4	3.1	-2.79	0.007
Learning disability*	3.6	2.4	-4.69	0
Mental retardation*	5.9	4.4	-3.62	0.001
Orthopedic impairment*	5.8	4.0	-2.06	0.044
Other health impairment	4.9	3.9	-1.43	0.158
Speech or language impairment*	2.2	1.8	-4.33	0
Low-incidence disability*	5.7	4.4	-2.15	0.036

*t, $p < .05$

SOURCE: U.S. Department of Education, National Center for Special Education Research, Pre-Elementary Education Longitudinal Study (PEELS), "Elementary School Teacher Questionnaire," "Kindergarten Teacher Questionnaire," and "Early Childhood Teacher Questionnaire."

setting outside the classroom, and home instruction remained fairly constant from 2003-04 to 2004-05 (see table 21).

Parent Satisfaction With Special Education Services

Parents were asked how satisfied they were with special education services their children received. The number of parents who reported being *very satisfied* decreased significantly ($t = -2.382$, $p = .02$). Forty-seven percent of parents reported being *very satisfied* in 2003-04 compared to 42 percent in 2004-05 (*S.E.* = 1.9 and 2.1, respectively) (see table 22). The number of parents who reported being *satisfied*, *dissatisfied*, and *very dissatisfied* did not change significantly. Forty-six percent of parents reported being *satisfied* in 2003-04 and 48 percent *satisfied* in 2004-05 (*S.E.* = 1.7 and 1.7, respectively) ($t = 1.314$, $p = .194$); 7 percent of parents reported being *dissatisfied* in both years (*S.E.* = 0.7 and 0.8, respectively) ($t = 1.013$, $p = .315$); and 2 percent of parents reported being *very dissatisfied* in both years (*S.E.* = 0.3 and 0.3, respectively) ($t = 1.391$, $p = .419$). So despite the other changes in educational services from 2003-04 to 2004-05, including time in various settings and types and number of services provided, parent satisfaction remained relatively stable, with a statistically significant decline only in the percentage of parents who were very satisfied.

Table 21. Mean hours per week that young children who received preschool special education services spent in various educational settings, by age cohort and school year: School years 2003-04 and 2004-05

	2003-04	2004-05
Total		
Regular education classroom*	8.2	15.0
Special education setting*	8.0	6.2
Therapy setting	0.8	0.8
Non-special education setting outside the classroom for remedial or special assistance	0.2!	0.2
Home instruction	0.2	0.2
Cohort A		
Regular education classroom*	4.3	7.1
Special education setting	8.7	9.0
Therapy setting*	0.6	0.8
Non-special education setting outside the classroom for remedial or special assistance	0.1	0.1
Home instruction	0.2!	0.1
Cohort B		
Regular education classroom*	6.7	13.5
Special education setting*	8.4	6.6
Therapy setting	0.7	0.8
Non-special education setting outside the classroom for remedial or special assistance	0.1	0.1
Home instruction	0.2!	0.1!
Cohort C		
Regular education classroom*	11.4	20.5
Special education setting*	7.2	4.4
Therapy setting	0.9!	0.8
Non-special education setting outside the classroom for remedial or special assistance	0.3!	0.4!
Home instruction	0.1!	0.2!

! Interpret data with caution.

*t, $p < .05$

SOURCE: U.S. Department of Education, National Center for Special Education Research, Pre-Elementary Education Longitudinal Study (PEELS), "Elementary School Teacher Questionnaire," "Kindergarten Teacher Questionnaire," and "Early Childhood Teacher Questionnaire."

Table 22. The percentage of young children who received preschool special education services whose parents were satisfied with special education services to various degrees: School years 2003-04 and 2004-05

	2003-04	2004-05
Very satisfied*	46.6	42.1
Satisfied	45.5	48.4
Dissatisfied	6.5	7.4
Very dissatisfied	1.5[!]	2.1

! Interpret data with caution.
* t, $p < .05$
NOTE: Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Special Education Research, Pre-Elementary Education Longitudinal Study (PEELS), "Parent Interview."

Summary

Teachers indicated that 89 percent of the children in the 2003-04 school year and 86 percent in 2004-05 received speech or language therapy. Occupational therapy (32% in 2003-04 and 35% in 2004-05) and learning strategies/study skills assistance by a special educator (30% in 2003-04 and 20% in 2004-05) were also commonly reported services both years. From 2003-04 to 2004-05, there was a statistically significant decrease in the percentage of children receiving nine types of services. The only service showing a significant increase was help from a one-to-one paraeducator or assistant. Furthermore, the mean number of special education and related services provided to young children with disabilities decreased from 3.5 in 2003-04 to 2.8 in 2004-05. Parent satisfaction remained relatively stable over that time period. There was a significant decline only in the number of parents who were *very satisfied*.

Chapter 5: A Year of Growth: The Knowledge and Skills of Preschoolers With Disabilities

Overall, preschoolers with disabilities who participated in the PEELS direct assessment in 2003-04 performed within one standard deviation of the national mean on tests of emerging literacy, early math proficiency, social-behavior skills, and motor skills. Mean scores on selected assessments varied by disability, age cohort, race/ethnicity, household income, and gender (Markowitz et al. 2006). This chapter takes those earlier analyses one step further by exploring change in scale scores from 2003-04 to 2004-05.

The augmented sample was used for the analyses presented in this chapter.[22] All analyses in this chapter were longitudinal, except those using the Adaptive Behavior Assessment System – II (ABAS-II), which were cross-sectional. To examine changes in scale scores from 2003-04 to 2004-05 for each subgroup, *t* tests for dependent samples were conducted. No analyses were conducted to compare differences in the amount of growth across the subgroups. The Benjamini-Hochberg procedure (Benjamini and Hochberg 1995) was used to control the family-wise error rate and avoid making false positive conclusions for multiple comparison analyses.

Emerging Literacy

This section describes the 1-year growth in emerging literacy skills for children who received preschool special education services and subgroups of that population defined by age cohort, gender, and disability category. Results for two assessments are included:

- the Woodcock-Johnson III: Letter-Word Identification and
- an adapted version of the Peabody Picture Vocabulary Test (PPVT).

Test scores for the emerging literacy measures reviewed in this section are standardized, so the mean for the population at each age is 100, for example. Even though children receive the same test items from year to year, more correct items are required for a 5-year-old to receive a standard score of 100 than for a 3-year-old to receive the same standard score. Population means remain constant, although subgroup means may go up or down. Performance on the various literacy measures that were included in the PEELS direct assessment suggest that children with disabilities maintained or improved their norm-referenced standing from school year 2003-04 to 2004-05. [23]

Woodcock-Johnson III: Letter-Word

The Letter-Word Identification test measures the child's word identification skills, without requiring the child to know the meaning of any word. It has a population mean of 100 and a standard deviation of 15. In school year 2003-04, young children who received preschool special education services performed as well, on average, as their peers without disabilities on Letter-Word Identification—with an overall mean performance of 98.2 (*S.E.* = 0.5). In school year 2004-05, the mean overall performance increased significantly to 100.2 (*S.E.* = 0.5) ($t = 7.427$, $p < .001$, see table 23).

---

[22] For children in the Supplemental Sample (6 percent of the augmented sample), we used imputed Wave 1 data.

[23] National norms for the assessments reported in this chapter are based on general population data, not data specifically for children with disabilities. National norms are used as benchmarks in the analyses to compare the PEELS children's growth from Wave 1 to Wave 2.

Table 23. Mean performance of young children who received preschool special education services on Woodcock-Johnson III: Letter-Word Identification, by age cohort: School years 2003-04 and 2004-05

	Total*	Cohort A*	Cohort B*	Cohort C*
2003-04	98.2	101.5	98.3	96.6
2004-05	100.2	102.7	100.9	98.2

* t, $p < .05$

SOURCE: U.S. Department of Education, National Center for Special Education Research, Pre-Elementary Education Longitudinal Study (PEELS), "Woodcock-Johnson III: Letter-Word Identification."

Between 2003-04 and 2004-05, the mean performance on the Letter-Word Identification test increased significantly for each age cohort. Children in Cohort A had an increase in their performance, from 101.5 (*S.E.* = 1.1) in 2003-04 to 102.7 (*S.E.* = 0.8) in 2004-05 ($t = 2.334$, $p = .023$). The mean performance of children in Cohort B increased from 98.3 (*S.E.* = 0.8) to 100.9 (*S.E.* = 0.9) ($t = 6.239$, $p < .001$), and the mean performance of children in Cohort C increased from 96.6 (*S.E.* = 0.6) to 98.2 (*S.E.* = 0.7) (see table 23) ($t = 4.475$, $p < .001$).

Both males and females had a significant improvement in performance. The mean performance of males increased significantly, from 97.4 (*S.E.* = 0.6) in 2003-04 to 99.7 (*S.E.* = 0.5) in 2004-05 ($t = 8.119$, $p < .001$). The mean performance of females also increased significantly, from 99.9 (*S.E.* = 0.9) in 2003-04 to 101.2 (*S.E.* = 0.8) in 2004-05 ($t = 2.545$, $p = .013$).

Children identified as having a developmental delay and children identified as having a speech or language impairment had statistically significant improvements in performance on the Letter-Word Identification test from 2003-04 to 2004-05 (DD: $t = 4.008$, $p < .001$; SL: $t = 5.691$, $p < .001$) (see table 24).

Table 24. Mean performance of young children who received preschool special education services on Woodcock-Johnson III: Letter-Word Identification, by Wave 1 primary disability: School years 2003-04 and 2004-05

	AU	DD*	ED	LD	MR	OI	OHI	SLI*	LI
2003-04	109.0	94.1	95.7	95.2	87.3	96.3	96.4	99.8	94.8
2004-05	105.5	96.9	96.7	97.9	85.2	98.4	93.8	102.2	98.1

* t, $p < .05$

NOTE: AU = Autism; DD = Developmental delay; ED = Emotional disturbance; LD = Learning disability; MR = Mental retardation; OI = Orthopedic impairment; OHI = Other health impairment; SLI = Speech or language impairment; LI = Low incidence.

SOURCE: U.S. Department of Education, National Center for Special Education Research, Pre-Elementary Education Longitudinal Study (PEELS), "Woodcock-Johnson III: Letter-Word Identification."

PPVT-III

The direct assessment included a measure of receptive vocabulary using an adapted version of the PPVT-III. Receptive vocabulary also is referred to as listening vocabulary or oral vocabulary. It is considered a strong predictor of language acquisition and cognitive development and is a key component in emerging literacy. Overall, the PPVT-III performance of young children who received preschool

special education services did not change significantly between 2003-04 and 2004-05. In 2003-04, the mean performance was 89.9 (*S.E.* = 0.7), and in 2004-05, it was 89.6 (*S.E.* = 0.6; see table 25).

However, children in Cohort A had a statistically significant increase in performance on the PPVT-III (t = 3.406, p = .001). Their mean scale scores increased from 88.5 (*S.E.* = 0.7) in 2003-04 to 88.9 (*S.E.* = 0.6) in 2004-05 (see table 25).

Table 25. Mean performance of young children who received preschool special education services on Peabody Picture Vocabulary Test III, by age cohort: School years 2003-04 and 2004-05

	Total	Cohort A*	Cohort B	Cohort C
2003-04	89.9	88.5	89.3	91.0
2004-05	89.6	88.9	89.0	90.4

*t, $p < .05$

SOURCE: U.S. Department of Education, National Center for Special Education Research, Pre-Elementary Education Longitudinal Study (PEELS), "Peabody Picture Vocabulary Test III."

The mean performance of males and females on the PPVT-III did not change significantly between 2003-04 and 2004-05. For males, scores averaged 89.6 (*S.E.* = 0.7) in 2003-04 and 89.2 (*S.E.* = 0.7) in 2004-05. For females, mean scores were 90.4 and 90.5 (*S.E.* = 1.2 and 1.1, respectively).

Table 26 reveals few statistically significant changes in mean performance on the PPVT-III by disability category. Overall, only the mean performance of children identified as having a low-incidence disability increased significantly, from 85.2 (*S.E.* = 1.8) in 2003-04 to 90.0 (*S.E.* = 1.7) in 2004-05 (t = 2.676, p = .01).

Table 26. Mean performance of young children who received preschool special education services on Peabody Picture Vocabulary Test III, by Wave 1 primary disability: School years 2003-04 and 2004-05

	AU	DD	ED	LD	MR	OI	OHI	SLI	LI*
2003-04	86.0	85.1	92.5	85.4	70.6	83.4	87.1	93.1	85.2
2004-05	81.1	85.0	94.1	85.0	66.5	86.7	82.0	93.4	90.0

* t, $p < .05$

NOTE: AU = Autism; DD = Developmental delay; ED = Emotional disturbance; LD = Learning disability; MR = Mental retardation; OI = Orthopedic impairment; OHI = Other health impairment; SLI = Speech or language impairment; LI = Low incidence.

SOURCE: U.S. Department of Education, National Center for Special Education Research, Pre-Elementary Education Longitudinal Study (PEELS), "Peabody Picture Vocabulary Test III."

Early Math Proficiency

Results for two norm-referenced math proficiency measures are presented in this section:

- Woodcock-Johnson III: Applied Problems, and
- Woodcock-Johnson III: Quantitative Concepts, including both concepts and number series.

Similar to the emerging literacy measures reviewed in the previous section, test scores for the early math measures reviewed in this section are standardized, so the mean for the population at each age is 100, for example.

Overall, performance on the early math measures significantly improved for PEELS children suggesting that children with disabilities maintained or improved their norm-referenced standing from school year 2003-04 to 2004-05.

Woodcock-Johnson III: Applied Problems

The Applied Problems test is a measure of children's ability to analyze and solve practical math problems using simple counting, addition, or subtraction operations. In school year 2003-04, the mean overall performance of young children who received preschool special education services was 90.8 (*S.E.* = 0.7). In school year 2004-05, the mean overall performance was 91.9 (*S.E.* = 0.7), which was a statistically significant increase ($t = 4.556$, $p < .001$).

The performance of children in Cohort A on the Applied Problems test increased from 88.9 (*S.E.* = 1.1) in 2003-04 to 91.7 (*S.E.* = 0.9) in 2004-05 (see table 27). Similarly, the mean performance of children in Cohort C increased from 90.7 (*S.E.* = 1.1) in 2003-04 to 93.1 (*S.E.* = 1.1) in 2004-05. These increases in performance were both statistically significant (Cohort A: $t = 5.608$, $p > .001$; Cohort C: $t = 5.046$, $p < .001$). The average performance of children in Cohort B did not change significantly.

Table 27. Mean performance of young children who received preschool special education services on Woodcock-Johnson III: Applied Problems, by age cohort: School years 2003-04 and 2004-05

	Total*	Cohort A*	Cohort B	Cohort C*
2003-04	90.8	88.9	91.9	90.7
2004-05	91.9	91.7	90.8	93.1

* t, $p < .05$

SOURCE: U.S. Department of Education, National Center for Special Education Research, Pre-Elementary Education Longitudinal Study (PEELS), "Woodcock-Johnson III: Applied Problems."

The mean performance for males on the Applied Problems test was 90.2 (*S.E.* = 0.8) in 2003-04 and increased significantly to 91.8 (*S.E.* = 0.7) in 2004-05 ($t = 5.704$, $p < .001$). Females had the same mean performance of 92.2 in both 2003-04 (*S.E.* = 1.1) and 2004-05 (*S.E.* = 1.2, see table 28).

Table 28. Mean performance of young children who received preschool special education services on Woodcock-Johnson III: Applied Problems, by gender: School years 2003-04 and 2004-05

	Male*	Female
2003-04	90.2	92.2
2004-05	91.8	92.2

* t, $p < .05$

SOURCE: U.S. Department of Education, National Center for Special Education Research, Pre-Elementary Education Longitudinal Study (PEELS), "Woodcock-Johnson III: Applied Problems."

Changes in mean performance on the Applied Problems test from 2003-04 to school year 2004-05 varied by disability category. In 2003-04, children identified as having a developmental delay had a mean performance of 83.7 (*S.E.* = 1.2), which increased to 86.0 (*S.E.* = 0.9) in 2004-05. This increase was statistically significant ($t = 3.488$, $p = .001$). Similarly, the mean performance of children identified as having a learning disability increased significantly ($t = 4.257$, $p < .001$). Children identified as having a speech or language impairment also had statistically significant increases in their mean performance ($t = 3.517$, $p = .001$) (see table 29).

Table 29. Mean performance of young children who received preschool special education services on Woodcock-Johnson III: Applied Problems, by Wave 1 primary disability: School years 2003-04 and 2004-05

	AU	DD*	ED	LD*	MR	OI	OHI	SLI*	LI
2003-04	81.8	83.7	91.6	85.3	62.7	91.0	84.0	96.1	84.2
2004-05	78.8	86.0	90.2	91.1	59.8	90.1	78.1	97.8	86.4

* $t, p < .05$

NOTE: AU = Autism; DD = Developmental delay; ED = Emotional disturbance; LD = Learning disability; MR = Mental retardation; OI = Orthopedic impairment; OHI = Other health impairment; SLI = Speech or language impairment; LI = Low incidence.

SOURCE: U.S. Department of Education, National Center for Special Education Research, Pre-Elementary Education Longitudinal Study (PEELS), "Woodcock-Johnson III: Applied Problems."

Woodcock-Johnson III: Quantitative Concepts

The Quantitative Concepts test measures basic mathematical concepts, symbols, and vocabulary. The score on this subtest is derived from questions that test important relational concepts and questions about numbers and number patterns. Children in Cohort C, the only cohort for which both years of data are available, had a statistically significant increase in performance on the Quantitative Concepts test, from 90.9 (*S.E.* = 0.8) in 2003-04 to 93.9 (*S.E.* = 0.9) in 2004-05 ($t = 5.945$, $p < .001$) (see table 30).

Table 30. Mean performance of young children who received preschool special education services on Woodcock-Johnson III: Quantitative Concepts, by age cohort: School years 2003-04 and 2004-05

	Total*	Cohort A	Cohort B	Cohort C*
2003-04	90.9	†	†	90.9
2004-05	93.9	†	94.0	93.9

† Not applicable.

* $t, p < .05$

SOURCE: U.S. Department of Education, National Center for Special Education Research, Pre-Elementary Education Longitudinal Study (PEELS), "Woodcock-Johnson III: Quantitative Concepts."

Social Behavior

This section presents results on children's social behavior based on a series of teacher rating scales that include the PKBS Social Skills, Adaptive Behavior Assessment System II – Self-Care, and Adaptive Behavior Assessment System II – Self-Direction scales. Note that for most children, different

teachers rated the child's behavior in school years 2003-04 and 2004-05, which may represent a source of bias.

Preschool and Kindergarten Behavior Scales 2 – Social Skills

The Social Skills scale of the PKBS-2 assesses age-appropriate personal and interpersonal behaviors of preschool and early elementary-age children. Age-appropriate personal behaviors include such things as, "works or plays independently," "follows rules," and "accepts decisions made by adults." Age-appropriate interpersonal behaviors include such things as "is cooperative," "comforts other children who are upset," and "takes turns with toys and other objects." Teachers were asked to rate how frequently the identified child exhibited a series of skills or behaviors such as those noted above during the previous 3 months. The measurement scale consists of four points, labeled *never*, *rarely*, *sometimes*, and *often*. A higher rating on the Social Skills composite index indicates a higher level of social adjustment. Scores are standardized with a mean of 100 and a standard deviation of 15. Rating scales were often completed by different teachers in 2003-04 and 2004-05, introducing a potential source of bias.

On the Social Skills subscale, mean teacher ratings for young children who received special education services increased significantly from 92.9 (*S.E.* = 0.5) in school year 2003-04 to 96.0 (*S.E.* =0.6) in 2004-05 ($t = 6.287$, $p < .001$) (see table 31). Children in Cohort A had a significant increase in teacher ratings on the Social Skills subscale, from 84.7 (*S.E.* = 0.9) in 2003-04 to 93.7 (*S.E.* = 0.9) in 2004-05 ($t = 11.06$, $p < .001$). Children in Cohort B also had a statistically significant increase in their mean teacher rating, from 93.7 (*S.E.* = 1.0) in 2003-04 to 95.7 (*S.E.* = 0.7) in 2004-05 ($t = 2.068$, $p = .043$). Changes for children in Cohort C were not significant (2003-04, $M = 96.6$, *S.E.* = 1.2; 2004-05, $M = 97.6$, *S.E.* = 1.1).

Table 31. Mean teacher ratings of young children who received preschool special education services on the Social Skills subscale of the Preschool and Kindergarten Behavior Scale, by age cohort: School years 2003-04 and 2004-05

	Total*	Cohort A*	Cohort B*	Cohort C
2003-04	92.9	84.7	93.7	96.6
2004-05	96.0	93.7	95.7	97.6

* t, $p < .05$

SOURCE: U.S. Department of Education, National Center for Special Education Research, Pre-Elementary Education Longitudinal Study (PEELS), "Social Skills subscale of the Preschool and Kindergarten Behavior Scale."

Both males and females had a statistically significant increase in performance on the Social Skills subscale from 2003-04 to 2004-05. The mean teacher rating for males increased from 91.4 (*S.E.* = 0.6) in 2003-04 to 94.4 (*S.E.* = 0.6) in 2004-05 ($t = 4.876$, $p < .001$) (see table 32). The mean for females increased from 96.6 (*S.E.* = 1.6) in 2003-04 to 99.8 (*S.E.* = 1.2) in 2004-05 ($t = 3.566$, $p = .001$).

Few statistically significant changes in teacher ratings were found on the Social Skills subscale when analyzed by disability category. However, mean teacher ratings for children identified as having a developmental delay ($t = 2.921$, $p = .005$) and for children identified as having a speech or language impairment ($t = 3.963$, $p < .001$) increased significantly (see table 33).

Table 32. Mean teacher ratings of young children who received preschool special education services on the Social Skills subscale of the Preschool and Kindergarten Behavior Scale, by gender: School years 2003-04 and 2004-05

	Male*	Female*
2003-04	91.4	96.6
2004-05	94.4	99.8

* t, $p < .05$

SOURCE: U.S. Department of Education, National Center for Special Education Research, Pre-Elementary Education Longitudinal Study (PEELS), "Social Skills subscale of the Preschool and Kindergarten Behavior Scale."

Table 33. Mean teacher ratings of young children who received preschool special education services on the Social Skills subscale of the Preschool and Kindergarten Behavior Scale, by Wave 1 primary disability: School years 2003-04 and 2004-05

	AU	DD*	ED	LD	MR	OI	OHI	SLI*	LI
2003-04	73.7	89.3	89.9	90.3	69.2	100.2	90.4	99.7	85.5
2004-05	74.7	93.3	87.8	99.7	71.0	99.1	102.0	102.5	86.0

* t, $p < .05$

NOTE: AU = Autism; DD = Developmental delay; ED = Emotional disturbance; LD = Learning disability; MR = Mental retardation; OI = Orthopedic impairment; OHI = Other health impairment; SLI = Speech or language impairment; LI = Low incidence.

SOURCE: U.S. Department of Education, National Center for Special Education Research, Pre-Elementary Education Longitudinal Study (PEELS), "Social Skills subscale of the Preschool and Kindergarten Behavior Scale."

Adaptive Behavior Assessment System II (ABAS-II) – Self-Care and Self-Direction

As mentioned previously, the Teacher/Daycare Provider (TDP) Form of the ABAS-II was used for children in preschool or not in school and the Teacher Form (TF) for children in kindergarten or elementary school. The following analyses are based upon cross-sectional data of different samples of children for 2003-04 and 2004-05, reported by different teachers, using different forms. Therefore, the average change in performance between the two waves was not calculated for the ABAS-II subscales. Scaled scores are based on a mean of 10 and a standard deviation of 3.

The Self-Care scale of the ABAS-II includes items that measure the child's basic personal care skills, including eating, dressing, bathing, toileting, grooming, and hygiene. The TDP Form includes items such as "swallows liquids with no difficulty" "nurses, drinks, or eats willingly, with little encouragement" and "tells teacher/daycare provider or other adult when he/she needs to use the bathroom." The TF includes some of the same items, but also includes items that are more appropriate for an older child. These items include such things as "uses the school restroom alone," "uses a fork to eat solid food" and "opens a tab-top can, milk carton, or screw-top bottle." For each of the skill areas, teachers rated the frequency with which the identified child exhibited specific behaviors using a 4-point scale, with ratings of the child's abilities as *is not able, never when needed, sometimes when needed,* and *always when needed.*

In 2003-04, the mean teacher/daycare provider rating on the Self-Care subscale for young children who received preschool special education services was 8.9 (*S.E.* = 0.1), and in 2004-05, it was 8.3 (*S.E.* = 0.2). The mean teacher rating for kindergarteners was 8.0 (*S.E.* = 0.4) in 2003-04 and 8.5 (*S.E.*

= 0.2) in 2004-05. The Self-Direction subscale measures the child's skills in independence, self-control, and personal responsibility. The TDP Form includes such items as child "shows interest in a toy or other object by looking at it for a few seconds," "follows an adult's request to quiet down or behave," and "looks for misplaced toys or games until he/she finds them." The TF includes such items as the child "routinely arrives at school or class on time," "works independently and asks for help only when necessary," and "stops a fun activity, without complaints, when told that time is up." As with the Self-Care subtest, scaled scores are based on a mean of 10 and a standard deviation of 3.

In school year 2003-04, the mean teacher/daycare provider rating on the Self-Direction subscale was 9.6 (*S.E.* = 0.1). In school year 2004-05, the mean rating was 9.5 (*S.E.* = 0.1). The mean for kindergarteners was 8.3 (*S.E.* = 0.3) in 2003-04 and 8.5 (*S.E.* = 0.2) in 2004-05.

Motor Skills

To provide a measure of each child's motor skills, the Fine Motor and Gross Motor subscales from the Vineland Adaptive Behavior Scales Classroom Edition were included in the teacher questionnaires. The results presented in this section are based on longitudinal analyses conducted on the data for the augmented sample. In school year 2003-04, the mean Motor Skills scale score for children who received preschool special education services was 94.4 (*S.E.* = 0.9). In school year 2004-05, the mean was 96.2 (*S.E.* = 1.0), representing a statistically significant increase ($t = 3.134$, $p = .003$).

The ratings of children in Cohort A significantly increased from 92.8 (*S.E.* = 0.9) in 2003-04 to 95.2 (*S.E.* = 1.0) in 2004-05 ($t = 3.207$, $p = .002$) (see table 34). Similarly, the mean ratings of children in Cohort C significantly increased from 95.4 (*S.E.* = 1.6) in 2003-04 to 98.4 (*S.E.* = 1.6) in 2004-05 ($t = 2.696$, $p = .009$). On average, the ratings of children in Cohort B did not change.

Table 34. Mean teacher ratings of young children who received preschool special education services on the Vineland Adaptive Behavior Scale: Motor Skills Domain, by age cohort: School years 2003-04 and 2004-05

	Total*	Cohort A*	Cohort B	Cohort C*
2003-04	94.4	92.8	94.3	95.4
2004-05	96.2	95.2	94.4	98.4

* $t, p < .05$

SOURCE: U.S. Department of Education, National Center for Special Education Research, Pre-Elementary Education Longitudinal Study (PEELS), "Vineland Adaptive Behavior Scale: Motor Skills Domain."

The overall mean ratings of males on the Motor Skills test increased significantly from 93.9 (*S.E.* = 0.8) in 2003-04 to 95.8 (*S.E.* = 1.0) in 2004-05 ($t = 2.804$, $p = .007$). Ratings on the Motor Skills test also increased significantly for male children in Cohort A and Cohort C (Cohort A: $t = 2.25$, $p = .028$; Cohort C: $t = 2.347$, $p = .022$). Mean ratings for children identified as having a learning disability increased significantly on the Motor Skills scale, from 90.5 (*S.E.* = 3.1) in 2003-04 to 101.9 (*S.E.* = 5.1) in 2004-05 ($t = 2.223$, $p = .03$). Children identified as having a developmental delay also had statistically significant increases in their ratings from 89.8 (*S.E.* = 1.2) in 2003-04 to 91.4 (*S.E.* = 1.4) in 2004-05 ($t = 2.722$, $p = .008$).

Alternate Assessment Results

Alternate assessments were completed for PEELS children who were not capable of participating in the direct assessment. For the alternate assessment, the child's teacher or service provider completed the Adaptive Behavior Assessment System (ABAS-II). An alternate assessment was completed for 331 children (12 percent of the total sample) in Wave 1 and 228 children (7 percent of the total sample) in Wave 2. Children who had an alternate assessment completed for them in Wave 1 did not necessarily have one completed for them in Wave 2, and vice versa. Thus, the following analyses are based on cross-sectional data of different samples of children for 2003-04 and 2004-05.

The children were identified for an alternate assessment in several ways. The first was during a screening interview prior to the assessment. If the respondent, who was familiar with the child, indicated that the child (1) could not understand and follow simple instructions that were spoken aloud or given in sign language; (2) had a visual impairment that would prohibit participation in an assessment that primarily involved pictures, text, and numbers; or (3) did not speak English or Spanish, then an alternate assessment was given. Alternate assessments were also completed for children who scored four or less on the two English-language subtests that were included in the direct assessment.

The ABAS-II is designed to evaluate whether an individual displays the adaptive skills necessary for daily living without the assistance of others. The adaptive skills measured by the ABAS-II are defined as "those practical, everyday skills required to function and meet environmental demands, including effectively and independently taking care of oneself and interacting with other people" (Harrison and Oakland 2003, p.3). Examples of adaptive skills that preschoolers use on a daily basis include those related to eating, dressing, expressing needs, interacting with peers, controlling one's behavior in a structured setting, and communicating with other people.

As noted in chapter 2, two forms of the ABAS-II were used for the PEELS alternate assessment. The TDP Form is designed for children ages 2 to 5, while the TF is designed for children in grades kindergarten through 12 (or ages 5 to 21). For each of the skill areas, teachers rated the frequency with which the identified child exhibited specific behaviors using a 4-point scale, with ratings of the child's abilities as *is not able, never when needed, sometimes when needed,* and *always when needed.* The specific skill areas measured by the ABAS-II used for the alternate assessment are Functional Academics, Self-Care, and Self-Direction, which were completed for all PEELS children as part of the teacher questionnaire, plus Communication, Community Use, School Living, Health and Safety, Leisure, Social, and Motor skills.[24] Data for children whose teachers completed the alternate assessment were not included in the Self-Care and Self-Direction analyses reported earlier in this chapter. The norm-referenced standardized scores for the ABAS-II have a mean of 10 and a standard deviation of 3.

Overall, the mean performance of preschoolers for whom an alternate assessment was completed was more than one standard deviation lower than the population mean for each of the skill areas measured by the ABAS-II. In most cases, the performance of preschoolers for whom an alternate assessment was completed was more than two standard deviations below the normed mean of 10. Table 35 presents the average teacher ratings (using the TDP Form) for PEELS children not yet in kindergarten on each of the ABAS-II skill areas. Table 36 presents the average teacher ratings (using the TF) for PEELS children in kindergarten or elementary school.

[24] The Community Use subscale is not included on the TDP Form. The Motor subscale is not included on the TF.

In 2003-04, mean teacher/daycare provider ratings for non-kindergarteners was 4.0 for the Motor subscale (*S.E.* = 0.2) and 2.2 for the Communication (*S.E.* = 0.2) and Social (*S.E.* = 0.1) subscales (see table 35). In school year 2004-05, non-kindergarteners' rating was 4.1 on the Self-Direction subscale (*S.E.* = 0.3) and 1.8 on the Communication subscale (*S.E.* = 0.1) (see table 35).

Table 35. Mean teacher ratings of young children who received preschool special education services and participated in an alternate assessment: ABAS-II—Skill area scores (Teacher/Daycare Provider Form): School years 2003-04 and 2004-05

	Total
Communication	
2003-04	2.2
2004-05	1.8
Functional (Pre) academics	
2003-04	3.2
2004-05	3.1
Health and safety	
2003-04	2.5
2004-05	2.6
Leisure	
2003-04	2.9
2004-05	3.0
Motor	
2003-04	4.0
2004-05	3.6
School living	
2003-04	3.5
2004-05	3.9
Self-care	
2003-04	3.2
2004-05	3.4
Self-direction	
2003-04	3.6
2004-05	4.1
Social	
2003-04	2.2
2004-05	2.3

SOURCE: U.S. Department of Education, National Center for Special Education Research, Pre-Elementary Education Longitudinal Study (PEELS), "ABAS-II—skill area scores."

In 2003-04, mean teacher ratings using the TF were 1.3 for the Communication (*S.E.* = 0.2) and Self-Direction (*S.E.* = 0.1) subscales and 3.8 for the Functional Academics (*S.E.* = 0.6) subscale. In school year 2004-05, mean teacher ratings using the TF were 1.4 for the Self-Direction subscale (*S.E.* = 0.1) and 3.8 for the Functional Academics (*S.E.* = 0.3) subscale (see table 36).

Table 36. Mean teacher ratings of young children who received preschool special education services and participated in an alternate assessment: ABAS-II—Skill area scores (Teacher Form): School years 2003-04 and 2004-05

	Total
Communication	
2003-04	1.3
2004-05	1.5
Community use	
2003-04	1.8
2004-05	1.8
Functional academics	
2003-04	3.8
2004-05	3.8
Health and safety	
2003-04	2.3[!]
2004-05	1.8
Leisure	
2003-04	2.1[!]
2004-05	2.6
School living	
2003-04	3.3
2004-05	2.9
Self-care	
2003-04	2.6[!]
2004-05	2.0
Self-direction	
2003-04	1.3
2004-05	1.4
Social	
2003-04	1.7
2004-05	1.8

! Interpret data with caution.

SOURCE: U.S. Department of Education, National Center for Special Education Research, Pre-Elementary Education Longitudinal Study (PEELS), "ABAS-II—skill area scores."

Summary

Data from 2004-05 indicate that children who received preschool special education services and took the direct PEELS assessment performed close to the mean for the norm-population on letter-word identification, social skills, and motor skills, and these scores were significantly higher than in 2003-04. The children scored within one standard deviation of the norm-referenced mean on a test of vocabulary and on both early math tests. Math scores were significantly higher in 2004-05 than in 2003-04. Children

who participated in the alternate assessment had mean ratings more than two standard deviations below the population mean in most skill areas.

References

Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of Royal Statistical Society*, *Series B*, *57*: 289-300.

Carlson, E. (2004). *Report of PEELS Sample Recruitment*. Rockville, MD: Westat.

Carlson, E. (1997). Outcomes for Students Declassified From Special Education. *Dissertation Abstracts International, 58*(11): 4231 (UMI No. 9815240).

Carlson, E., and Parshall, L. (1996). Academic, Social, and Behavioral Adjustment for Students Declassified From Special Education. *Exceptional Children, 63*: 89-100.

Carrow-Woolfolk, E. (1995). Oral and Written Language Scales: Listening Comprehension and Oral Expression. Circle Pines, MN: American Guidance Service.

Connors, C.K. (1990). *Connors' Rating Scales Manual*. North Tonawanda, NY: Multi-Health Systems.

Council for Exceptional Children. (2000). *Making Assessment Accommodations: A Toolkit for Educators.* Reston, VA: Author.

Danaher, J., Kraus, R., Armijo, C., and Hipps, C. (2003). *Section 619 Profile*. Chapel Hill, NC: National Early Childhood Technical Assistance Center.

De Avila, E., and Duncan, S. (1998). *preLAS 2000.* Monterey, CA: CTB McGraw-Hill.

Dunn, L.M., and Dunn, L.M. (1997). *Peabody Picture Vocabulary Test-Third Edition*. Circle Pines, MN: American Guidance Services.

ECRI-MGD (Early Childhood Research Institute on Measuring Growth and Development). (2001). *Individual Growth and Development Indicators*. Minneapolis, MN: Center for Early Education and Development, University of Minnesota.

Gresham, F.M., and Elliott, S.N. (1990). *Social Skills Rating System*. Circle Pines, MN: American Guidance Service.

Halgren, D.P., and Clarizio, H.F. (1993). Categorical and Programming Changes in Special Education Services. *Exceptional Children, 59*: 547-555.

Harrison, J., and Oakland, T. (2003). *Adaptive Behavior Assessment System* (2nd edition). San Antonio, TX: Harcourt Assessment Company, The Psychological Corporation.

Individuals with Disabilities Education Act, P. L. 05-17, 105 Stat. 587 (codified as amended at 20 U.S.C. §1400 et seq.

Innocenti, M.S. (2005). *Classification Status of Children Once Enrolled in Early Intervention Programs*. Report submitted to the Utah State Office of Education.

Kaufman, A.S., and Kaufman, N.L. (1993). *Kaufman Adolescent and Adult Intelligence Test*. Circle Pines, MN: American Guidance Service.

Kaufman, A.S., and Kaufman, N.L. (1990). *Kaufman Brief Intelligence Test*. Circle Pines, MN: American Guidance Service.

Kaufman, A.S., and Kaufman, N.L. (1985). *Kaufman Test of Educational Achievement*. Circle Pines, MN: American Guidance Service.

Lambert, N., and Windmiller, M. (1981). *AAMD Adaptive Behavior Scale – School Edition*. Monterey, CA: Publishers Test Service.

Markowitz, J., and Bridges-Cline, F. (1991, November). *A Six-Year Study of Children Enrolled in Special Education Under the Age of Five Years*. Rockville, MD: Montgomery County Public Schools.

Markowitz, J., Carlson, E., Frey, W., Riley, J., Shimshak, A., Heinzen, H., Strohl, J., Lee, H., and Klein, S. (2006). *Preschoolers' Characteristics, Services, and Results: Wave 1 Overview Report from the Pre-Elementary Education Longitudinal Study (PEELS)*. Rockville, MD: Westat.

Markwardt, F. (1989). *Peabody Individual Achievement Test-Revised*. Circle Pines. MN: American Guidance Service.

Matson, J.L., Esvelt-Dawson, K., and Kazdin, A.E. (1983). Validation Methods for Assessing Social Skills in Children. *Journal of Clinical Psychology, 12*: 175-180.

McGrew, K.S., and Woodcock, R.W. (2001). Technical Manual. *Woodcock-Johnson III.* Itasca, IL: Riverside Publishing.

Merrell, K.W. (2002). *Preschool and Kindergarten Behavior Scales, Second Edition.* Austin, TX: PRO-ED.

Merrell, K.W. (1995). Relationships Among Early Childhood Behavior Rating Scales: Convergent and Discriminant Construct Validity of the Preschool and Kindergarten Behavior Scales. *Early Education and Development, 6*: 253-264.

Merrell, K.W. (1993). *School Social Behavior Scales*. Eugene, OR: Assessment-Intervention Resources.

Roid, G.H., and Miller, L.J. (1995, 1997). *Leiter International Performance Scale-Revised*. Wood Dale, IL: Stoelting Co.

Sparrow, S.S., Balla, D.A., and Cicchetti, D.V. (1985). *Vineland Adaptive Behavior Scales* (Classroom Edition), Circle Pines, MN: American Guidance Service.

Stocking, M.L., and Lord, F.M. (1983). Developing a Common Metric in Item Response Theory. *Applied Measurement, 7*: 201-210.

Thurlow, M.L., Lehr, C.A., and Ysseldyke, J.E. (1987). Exit Criteria in Early Childhood Programs for Handicapped Children. *Journal of the Division for Early Childhood, 11*(2): 118-123.

Thurlow, M.L., Ysseldyke, J.E., and Weiss, J.A. (1988). Early Childhood Special Education Exit Decisions: How Are They Made? How Are They Evaluated? *Journal of the Division for Early Childhood, 12*(3): 253-262.

U.S. Department of Health and Human Services (US HHS), Administration for Children and Families, Head Start Bureau. (2005, February). *Report to Congress: Assessment of Children in Head Start Programs.* Washington, DC: U.S. Government Printing Office.

VESID Program Quality and Its Effect on the Placement of Preschool Education Children in School-age Programs. Special Education Quality Indicator Study. August 2003.

Walker, D., Singer, J., Palfrey, J., Orza, M., Wenger, M., and Butler, J. (1988). Who Leaves and Who Stays in Special Education: A 2-Year Follow-Up Study. *Exceptional Children, 54*: 393-402.

Wechsler, D. (1992). *Wechsler Individual Achievement Test.* San Antonio, TX: Psychological Corporation.

Wechsler, D. (1991). *Wechsler Intelligence Scale for Children–Third Edition.* San Antonio, TX: Psychological Corporation.

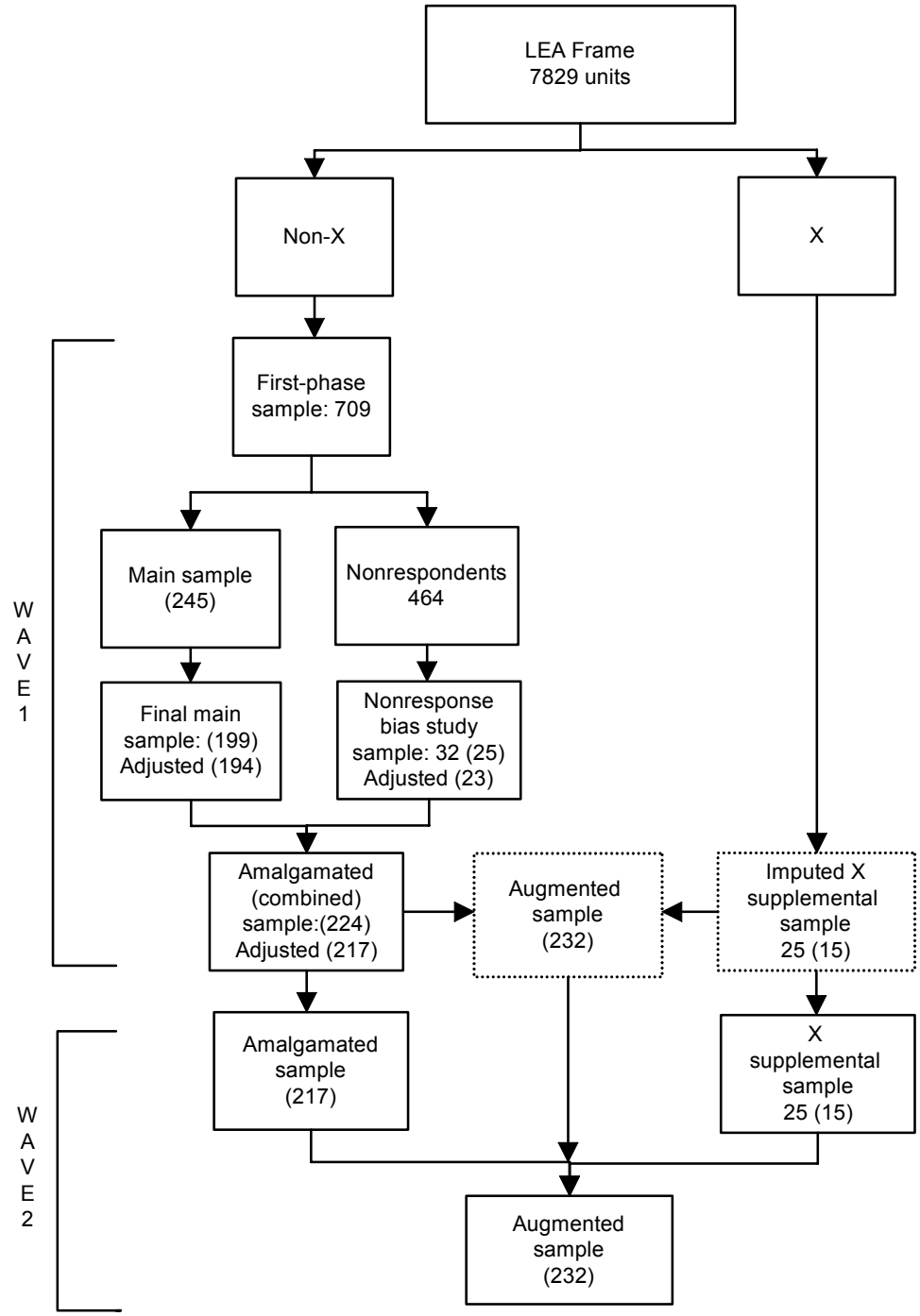
Westat. (2002). *WesVar Version 4.2*. Rockville, MD: Author.

Wong, M.M. (1997). Patterns of Special Education Placement for Preschool and School-age Children. *Dissertation Abstracts International, 58*(12): 4616 (UMI No. 9819362).

Woodcock, R.W., McGrew, K.S., and Mather, N. (2001). *Woodcock-Johnson III Tests of Achievement*. Itasca, IL: Riverside Publishing.

Woodcock, R. and Muñoz-Sandoval, A. (1996). *Pruedas de Aprovechamiento Revisada*. Itasca, IL: Riverside Publishing.

Appendix A: Diagram of Selection of LEA Sample



Note: X stands for the state that originally did not participate. LEA counts for X and non-X were suppressed for confidentiality reasons. The figures in parentheses are the number of participating LEAs. They were adjusted as the LEAs, which did not contribute any data, were dropped. The dotted boxes represent a mirror image created by imputation of the X supplemental sample selected in Wave 2.

Appendix B: Weighting Procedures

This appendix describes weighting procedures used in Waves 1 and 2 of PEELS. The PEELS study was designed to use a nationally representative sample of local education agencies (LEAs) and children ages 3 through 5 with disabilities to generate weighted estimates that reflect the characteristics of the population, not the sample.

District Weighting

The LEA weighting procedure includes developing base weights and replicate weights. Replicate weights were generated for each set of full-sample weights to allow the creation of estimated standard errors on all statistics.

District Base Weights

Calculation of the base weights started with the first-stage sample of 709 LEAs for the amalgamated sample and 25 LEAs for the supplemental sample. Analysis of nonresponse patterns revealed that nonresponse adjustment to the base sampling weights for the main sample could be carried out within the design stratum cells. Therefore, district base weights were recomputed within each sampling stratum cell as the number of districts on the sampling frame divided by the number of districts that participated in the study. The sum of the base weights represents 7,829 districts.[25] These weights will be denoted as w_h, which is the same for all LEAs within a stratum cell (defined by district size, region, and wealth category for non-supplemental LEAs and by district size alone for supplemental sample LEAs).

Replicate Weights

Replicate weights were developed to facilitate variance estimation using Westat's proprietary software, WesVar. Due to restrictions in the DAS software that will be used for data dissemination, the jackknife method JK2 with 62 replicates was used instead of the JKn method used previously for Wave 1 weighting.

The JK2 method requires defining the variance strata and two variance units per variance stratum. The variance strata were defined by the sampling strata by size, region, and wealth at the beginning. However, sampling strata with no or a small number of responding LEAs were collapsed with a neighboring stratum cell with similar sampling rates. Sampling strata with a large number of LEAs were split into two variance strata. Altogether, 62 variance strata were created. Variance units were formed by randomly grouping districts within each variance stratum up to three variance units. The number of groups was determined by the number of replicates.

The replicate weights were then created for the JK2 method. If there are two variance units, this is done by assigning a zero weight to records in one variance unit chosen randomly and doubling the weights for records in the other variance units from the same variance stratum but leaving the weights for records in other variance strata unchanged. If the randomly chosen variance unit from the *i*-th variance

[25] This number is different from the total number of LEAs in the country because the smallest LEAs were not covered by the sample design.

stratum is denoted as U_{i1} and the other variance unit as U_{i2}, algebraically the *i*-th replicate weight for the *j*-th LEA record, w_{ij}^{*}, is given by

$$w_{ij}^{*} = \begin{cases} 0 & \text{if the } j \text{ - th record is in } U_{i1} \\ 2w_h & \text{if the } j \text{ - th record is in } U_{i2} \\ w_h & \text{if the } j \text{ - th record is not in the } i \text{ - th variance stratum} \end{cases}$$

where w_h is the full sample base weight for the stratum cell *h* to which the *j*-th LEA belongs, *i* = 1, 2, …, 62; *j* = 1, 2, …, 232.

If there are three variance units, replicate weight calculation is more complex. In this case, another variance stratum number is needed; usually, an existing number is arbitrarily assigned. Let this be *k* and the three variance units be randomly ordered as U_{i1}, U_{i2}, and U_{i3}. The replicate weight that corresponds to this situation is defined as:

$$w_{ij}^{*} = \begin{cases} 0 & \text{if } j \text{ - th record is in } U_{i1} \\ 1.5w_h & \text{if } j \text{ - th record is in } U_{i2} \\ 1.5w_h & \text{if } j \text{ - th record is in } U_{\mathrm{i3}} \end{cases}$$

and

$$w_{kj}^{*} = \begin{cases} 1.5w_h & \text{if } j \text{ - th record is in } U_{i1} \\ 0 & \text{if } j \text{ - th record is in } U_{i2} \\ 1.5w_h & \text{if } j \text{ - th record is in } U_{\mathrm{i3}} \end{cases}$$

Consequently, each LEA has a base weight w_h and 62 replicate weights, w_{1j}^{*}, w_{2j}^{*}, …, w_{62j}^{*}.

Child Weighting: Within LEA Child Base Weight

After the child sampling was finished, the sampling status was defined by child status ID, which has 15 categories shown in table B-1.

The status codes 1, 2, and 4 are interim codes, and no child should have this code at the end of data collection in each wave. A large number of children have a status code of 3 since they were passed through the sampling system but not selected into the sample (those who were selected had a code value of 4 but subsequently moved to one of the remaining categories). Only children in category 6 are enrolled for the study. Children in categories 9 and 11 were selected first but then deselected due to the maximum 80-children limit for each district or district-wide non-participation. These and 1, 2, 8, and 12 are treated as not passed in the sampling system. Status codes 60, 61, and 62 are relevant only to the children in Wave 2.

Table B-1. Child status codes

Code	Definition	Description
1	Entering	The child record is entered into the computer system.
2	Ready sample	The child record is ready for sampling.
3	Sampled	The child record has gone through the sampling system.
4	Selected	The child record is selected into the sample.
5	Ineligible	The child is ineligible.
6	Enrolled	The child is enrolled for the study.
7	Declined	The child has declined.
8	Max reached/not sampled	The record is not sampled because the district has reached the cap of 80.
9	Max reached/deselected	The record is selected but subsequently deselected because the district has reached the cap of 80.
10	Nonresponse	The child was selected but did not respond.
11	Deselected-No LEA/child participation	The child was selected but subsequently deselected because neither LEA questionnaire was filled out nor any child participated in the study.
12	Desampled/district nonparticipation	The child was sampled but subsequently desampled because the whole district dropped out of the study.
60	Deceased	The child died after Wave 1.
61	Ineligible	The child turned out to be ineligible after Wave 1.
62	Study withdrawal	The child withdrew from the study after Wave 1.

Child sampling was done using the sampling system within sampling strata (called LEA-cohort) defined by District ID and the five cohort IDs [3-years-old ongoing (A_O), 4-years-old ongoing (B_O), 4-years-old historical (B_H), 5-years-old ongoing (C_O), 5-years-old historical (C_H)].

During reweighting it was found that nine children had incorrect birthdates. The correction of their birthdates altered their sampling LEA-cohort strata. We recomputed sampling rates of those affected LEA-cohort strata assuming the realized strata are the real strata from which they were selected. Four children from two LEAs swapped their LEA-cohort strata within their LEAs, and thus no change in the sampling rate was necessary for them. This approach may be termed as conditional on the realized LEA-cohort strata. This may introduce some bias but will reduce the variance. We believe that the bias introduced by this approach is negligible because the number of problem cases is small, and the sampling rate changes are not great.

A within-LEA base sampling weight for children by child sampling stratum was created for all sampled and selected children (categories 5, 6, 7, 10, 60, 61, 62) based on the sampling rate. The weight for a selected child i in an LEA-cohort within LEA stratum h is defined as the inverse of the sampling rate that was applied:

$$w_{hi}^{c} = \frac{1}{r_{hi}} .$$

Note that the subscript i now identifies sample children, so it has a different meaning from the one used in the previous section. The sampling rate r_{hi} depends on the LEA stratum h, where the child's LEA is contained, and the child's particular LEA-cohort.

The sampling rate changed during the sampling process for many LEA-cohort strata, so children in those LEA-cohort strata were selected with a different sampling rate from that of other children in the same LEA-cohort stratum, depending on the time of sampling. Therefore, the children from the same LEA may have different base weights.

The sum of unconditional base weights in a cohort is close but not equal to the child list total of the cohort. We first considered using a conditional approach that defines the within-LEA child weight based on the realized sample size instead of using the sampling rate. This approach cuts down the variance due to random sample sizes that resulted from the Bernoulli sampling procedure used for child sampling from the ongoing lists. However, this approach became problematic because 48 LEA-cohort strata did not have any children selected due to small sampling rates and inaccurate list size estimates used to calculate the sampling rates and also by chance. Therefore, if we used the conditional approach, children from the 48 LEA-cohort strata would not be represented. To avoid this problem, we used the unconditional approach and the corresponding formula given above.

There are two exceptions to using unconditional weights:

- First, for LEA-cohort strata that have some children in categories 1, 2, 8, and 9, we used the conditional weighting method because not all the children were covered by the unconditional weighting; that is, some children were unsampled or deselected, which makes the sampling rate used for sample selection wrong. For these cases, the conditional weight was calculated by dividing the child list total of the LEA-cohort by the actual number of children selected for the LEA-cohort:

$$w_{hi}^{c} = \frac{N_{hi}}{n_{hi}}.$$

  The conditional weight was the same for every child and summed exactly to the list total of the LEA-cohort stratum.

- Second, after we performed the weighting using the methods above, we checked the sum of weights against the list counts, by cohort, and found some large differences, which were mainly due to large discrepancies for the following LEA-cohorts: 1457B_O, 1457C_O, 3319C_H, 3495C_O, 1060C_O, 2044B_H, 2596B_H, 1917C_H, 1519B_H, 3256B_H, 9002A_O, 9002_B_O, 2549C_H, 1519A_O, 2864B_H, and 1472B_H. We recalculated the sampling weights using the conditional approach for them.

With this correction, the sum of weights was almost the same as the overall list total. The weights also agree quite well at various levels of aggregation.

Child Base Weight

The overall weight for the selected children was created by multiplying the child base weight and the LEA full sample weights, w_h, defined earlier:

$$w_{hi} = w_h w_{hi}^{c}.$$

The overall child replicate weights are then obtained by multiplying the child base weight and the LEA replicate weights.

Noncoverage Adjustment for Smallest LEAs

In the PEELS sample design, size 5 (very small) LEAs were not sampled. This is because size 5 LEAs accounted for only a small percentage of the whole target population but required more resources to sample because they are numerous. We decided to adjust for the noncoverage of size 5 children by increasing the size 4 children's base weights by a ratio factor calculated from the original frame stratified by region and wealth. Note that only size 4 children's weights are adjusted. The adjusted weights are given by

$$w_{hi}^{*} = \begin{cases} w_{hi}, & \text{if size less than 4,} \\ w_{hi} f_{hi}^{\text{cov}}, & \text{if size = 4,} \end{cases}$$

where f_{hi}^{cov} is the coverage adjustment factor for size 4 LEAs. Table B-2 shows the factors by region and wealth class.

Table B-2. Non-coverage adjustment factors

Region	Wealth	Non-coverage factor
1	1	1.0798
1	2	1.1203
1	3	1.2089
1	4	1.4796
2	1	1.0530
2	2	1.0391
2	3	1.0517
2	4	1.0699
3	1	1.1428
3	2	1.2300
3	3	1.4222
3	4	1.5694
4	1	1.2022
4	2	1.3007
4	3	1.3887
4	4	1.4203

Nonresponse Adjustment of Child Base Weight

The child base weights were adjusted to compensate for the nonresponding sample children. Each of the four input datasets contain all the children who have child status ID equal to 5, 6, 7, or 10, where 5 = ineligible, 6 = enrolled, 7 = declined, and 10 = nonresponse. Only children with child status ID = 6 are enrolled in the study. The eligibility of children with status 10 was unknown for most records; however, for 182 records this could be determined by a subcoded value of child status ID (see table B-3). The weights of the enrolled children were adjusted to account for the unknown eligibility and nonresponse.

Table B-3. Subcodes for child eligibility

Code	Description	Eligibility
1	Received, eligibility status not reported/not known	Unknown
2	Received, eligible case, district could not reach family	Known
3	Received, eligible case, problem not resolved	Known
4	Enrollment form not received	Unknown
5	Enrollment form received late	Unknown

We first tried to use CHAID analysis to define the adjustment cells for the main sample based on the size, region, wealth, age, and placement on the ongoing or historical lists. We found that the stratification variables size, region, and wealth were the most significant predictors of nonresponse. We decided to use the stratification cell as the initial nonresponse adjustment cell.

Since the eligibility of some children was not known, adjustment was done in two stages. First, the nonresponse status was redefined as

Status	Meaning
1	Enrolled
2	Eligible but declined
3	Ineligible
4	Nonresponse, eligibility unknown

In the first stage adjustment, the adjusted weight was $w_{hi}^{**} = w_{hi}^{*} f_{hi}^{NR1}$, where f_{hi}^{NR1} is the factor defined in the table below. S_j is defined as the sum of weights of all cases within each of the nonresponse cells. The nonresponse adjustment factor f_{hi}^{NR1} is then determined depending on the child sample status by:

Status	Adjustment factor
1	$\frac{S_1 + S_2 + S_3 + S_4}{S_1 + S_2 + S_3}$
2	$\frac{S_1 + S_2 + S_3 + S_4}{S_1 + S_2 + S_3}$
3	$\frac{S_1 + S_2 + S_3 + S_4}{S_1 + S_2 + S_3}$
4	0

In the second stage adjustment, the adjusted weight is $w_{hi}^{***} = w_{hi}^{**} f_{hi}^{NR2}$, where the nonresponse adjustment factor f_{hi}^{NR2} is determined as follows:

Status	Adjustment factor

1	$\frac{S_1 + S_2}{S_1}$
2	0
3	1

Truncation of Weight Outliers for Child Base Weights

After nonresponse adjustment, we truncated the weight outliers within five cohorts (A_O, B_O, B_H, C_O, and C_H). This was deemed necessary because the weights vary too much to contain the variance at a reasonable level. Sometimes a simple rule, such as the three-median rule, was used to set truncation of boundary. This rule truncates weights that are larger than three times the median weight to three times the median weight:

$$w_{hi}^{***} = \begin{cases} w_{hi}^{***}, & \text{if } w_{hi}^{***} \le 3\text{Median}, \\ 3\text{Median}, & \text{if } w_{hi}^{***} > 3\text{Median}. \end{cases}$$

However, for some child sampling strata, the three-median rule caused too many weights to be truncated. We tried to keep the percentage of truncated weights to less than 3 percent, so, for some child sampling strata, we used a three-and-a-half-median or four-median rule. For the children who had their full sample weight truncated, all the replicate weights were reduced by the same percentage.

Post-stratification of Enrolled Child Weight

The nonresponse adjusted children's weight was further adjusted by a post-stratification procedure. The control totals for post-stratification contained the number of special education children enrolled by December 2003, by age, for each of the 50 states and the District of Columbia.

Post-stratification was necessary because several states did not have any children sampled, either because, by chance, no LEAs in those states were selected, or none of the selected LEAs in a state responded. It should be noted that the control totals are snapshot figures, while the PEELS population includes children enrolled during a certain time period. The control totals also include children from the very small (size 5) school districts, which were not covered (but were adjusted for) by the PEELS sample.

The post-strata were formed by crossing the three age groups and nine subregions formed by combining states within the same region by their geographical proximity. The size of states in terms of number of children was also taken into consideration in order to obtain similar-sized post-strata.

After the post-stratification was applied, we created the final enrolled children's base weight. This weight is called the children's base weight, although it resulted from various adjustments, because it will be the base for further nonresponse adjustments for different data collection instruments. These are discussed in the following section.

Parent Interview Weights

The parent interview was attempted for all enrolled children, but some parents did not respond. The weights for the parent interview data were created by adjusting the enrolled children's base weights for parent nonresponse. The nonresponse adjustment cells were the same as the ones formed for the nonresponse adjustment to obtain the enrolled children's base weight. This worked well because the response rate for the parent interview was very high. In Waves 1 and 2 at the completion of imputation, parent interview data and corresponding weights were available for 96 percent and 93 percent, respectively, of the children in the augmented sample. Parent interview data and corresponding weights were available for 91 percent of the children in the augmented sample in both waves.

Child Assessment Weights

The child assessment was done in two ways. Most of the children were assessed directly, but for children who could not complete the direct assessment, an alternate assessment was conducted. Together, they represent the whole population of either directly assessable children or unassessable children. The child assessment weight was created by using the enrolled children's weights as base weights and adjusting for child nonresponse in the assessment data. The nonresponse adjustment cells were the same as the ones formed for the nonresponse adjustment to create the enrolled children's base weight. The response rate for child assessment was very high. In Waves 1 and 2 at the completion of imputation, assessment data and corresponding weights were available for 96 percent and 95 percent, respectively, of the children in the augmented sample. Assessment data and corresponding weights were available for 92 percent of the children in the augmented sample in both waves.

Teacher Weights

The teacher interview was attempted for the teachers of all enrolled children, but some teachers did not respond. The weights for the teacher interview data were created by adjusting the enrolled children's base weights for teacher nonresponse. The nonresponse adjustment cells were the same as the ones formed for the nonresponse adjustment to create the enrolled children's base weight. The response rate for teachers was lower than for parents and child assessment. In Waves 1 and 2 at the completion of imputation, teacher interview data and corresponding weights were available for 79 percent and 84 percent, respectively, of the children in the augmented sample. Teacher interview data and corresponding weights were available for 65 percent of the children in the augmented sample in both waves.

Parent-Child Weights

In many analyses, both parent interview and child assessment information are needed; the parent-child weight was for children with both child assessment data and parent interview data. The enrolled children's weights were used as base weights and adjusted for the nonresponse of children in the parent-child data. The nonresponse cells were the same as the ones formed in the nonresponse adjustment for children's base weight. In Waves 1 and 2 at the completion of imputation, both parent interview and assessment data and corresponding weights were available for 92 percent and 89 percent, respectively, of the children in the augmented sample. Child assessment and parent interview data and corresponding weights were available for 85 percent of the children in the augmented sample in both waves.

Parent-Child-Teacher Weights

In some analyses, information from all three instruments is needed. The parent-child-teacher weight is for children with completed interviews for parent interview, child assessment, and the teacher interview. The enrolled children's weights were used as base weights and adjusted for the nonresponse of children in the parent-child data. The nonresponse cells were the same as the ones formed in the nonresponse adjustment for children's base weight. Because of the lower response rate in the teacher interview, the response rate for the parent-child-teacher data is relatively low. In Waves 1 and 2 at the completion of imputation, child assessment, parent interview, and teacher interview data and corresponding weights were available for 70 percent and 76 percent, respectively, of the children in the augmented sample. Child assessment, parent interview, and teacher interview data and corresponding weights were available for 57 percent of the children in the augmented sample in both waves.

Use of Weights in Analysis

Table B-4 provides a description of each weight available after Wave 2 and the analyses for which it is used. For this report, cross-tabulations with covariates from the PEELS demographics file, such as age cohort, sex, and race/ethnicity, use Wave 1 and Wave 2 cross-sectional weights because the demographics file has no missing data and no specific weights. Cross-tabulations with covariates from the Wave 1 files, such as household income, use Wave 1 cross-sectional weights for the Wave 1 cross-tabulations and longitudinal weights for the Wave 2 cross-tabulations because the Wave 2 cross-tabulations use data from Wave 1 and Wave 2 sources. Wave 2 cross-sectional weights were used in table columns with Wave 2 covariates or demographics analyzed with Wave 2 dependent variables.

Table B-4. Description and uses of Wave 1 and Wave 2 cross-source and longitudinal weight variables used in this report

Description	Use of weight
Cross-sectional Wave 1 assessment weight	Analyses using only data from the Wave 1 assessment file
Cross-sectional Wave 2 assessment weight	Analyses using only data from the Wave 2 assessment file
Longitudinal assessment weight	Analyses using only data from the Wave 1 and Wave 2 assessment files
Cross-sectional Wave 1 parent interview weight	Analyses using only data from the Wave 1 parent interview file
Cross-sectional Wave 2 parent interview weight	Analyses using only data from the Wave 2 parent interview file
Longitudinal parent interview weight	Analyses using only data from the Wave 1 and Wave 2 parent interview files
Cross-sectional Wave 1 teacher weight	Analyses using only data from the Wave 1 teacher files
Cross-sectional Wave 2 teacher weight	Analyses using only data from the Wave 2 teacher files
Longitudinal teacher weight	Analyses using only data from the Wave 1 and Wave 2 teacher files
Cross-sectional Wave 1 parent/assessment weight	Analyses using data from the Wave 1 parent interview and Wave 1 assessment files
Cross-sectional Wave 2 parent/assessment weight	Analyses using data from the Wave 2 parent interview and Wave 2 assessment files
Cross-sectional Wave 2 parent/assessment/ teacher weight	Analyses using data from the Wave 2 parent interview, Wave 2 assessment, and Wave 2 teacher files
Longitudinal parent/assessment/teacher weight	Analyses using Wave 1 and Wave 2 data from (1) the parent interview, (2) assessment, or (3) teacher files and Wave 1, Wave 2, or Wave 1 and Wave 2 data from the other types of files (parent interview, assessment, or teacher)
Longitudinal parent/assessment weight	Analyses using Wave 1 and Wave 2 data from the parent interview or assessment files and (1) Wave 1, (2) Wave 2, or (3) Wave 1 and Wave 2 data from the other type of file (assessment or parent interview)
Cross-sectional Wave 1 parent/assessment/ teacher weight	Analyses using data from the Wave 1 parent interview, Wave 1 assessment, and Wave 1 teacher files

Appendix C: Results from PEELS Nonresponse Bias Study

This report presents results of a nonresponse bias analysis of PEELS Wave 1 data. The study was conducted in response to concerns about potential bias from low stage 1 response rates. As a result, terms of clearance for the Pre-Elementary Education Longitudinal Study (PEELS) (OMB #1820-0656) required the U.S. Department of Education's Office of Special Education (OSEP) to submit to the Office of Management and Budget (OMB) a nonresponse analysis report.

To provide the needed confidence to data users, data producers, and study sponsors, OSEP funded a small-scale sample survey of LEAs that initially did not agree to participate in PEELS (464 LEAs or 65 percent of the original LEA sample). Westat selected a random sample of 32 nonparticipating LEAs in Wave 1, allocating the sample to the existing size strata. While 25 of those LEAs agreed to participate, only 23 (72%) actually followed through with their participation, meaning they successfully recruited one or more families[26]. This nonresponse study sample is roughly 10 percent of the size of the main LEA sample. Table C-1 shows the size distribution of the LEAs participating in the nonresponse study.

Table C-1. Frequency of LEAs in PEELS by size stratum and sample type

Size stratum	U.S.	Main sample	Nonresponse sample
Total	7,818	194	23
Very Large	117	33	2
Large	629	32	5
Medium	1,897	43	6
Small	5,175	86	10

The instruments and data collection procedures were exactly the same for the main and nonresponse study participants, so any differences between the two samples can be attributed to the differences in the characteristics of the subpopulations that the samples represent (main study sample and nonresponse study sample).

This nonresponse bias study has three primary research questions. They include the following:

1. Can we produce weighted data from the main sample that provides unbiased national estimates of student performance on key outcome variables?

2. Do statistical differences exist between the performances of students in participating districts and students in nonresponse study districts on key outcome variables?

3. Is student performance on key outcome variables a factor in the decision to participate or not in PEELS?

[26] Nonresponse may cause some bias in estimates obtained from a sample of only respondents if nonrespondents are different from respondents in terms of their characteristics of interest. Nonresponse adjustment weighting was performed so that the bias due to nonresponse is minimized. Even if the nonresponse adjustment weighting was not perfect, bias would not be serious because the response rate of 72 percent is reasonably high.

Methods Used to Analyze Nonresponse Bias

Our general strategy for assessing bias due to nonresponse includes three types of analyses. The first set of analyses involves comparisons between weighted data of the *main* sample versus weighted data of the *amalgamated* sample (which includes the main and nonresponse samples). The second set of analyses compares unweighted data in the main sample with the nonresponse sample. A final set of analyses involves logistic regressions using participation status as the dependent variable and child performance among the independent variables. Each of these analyses is discussed in more detail below.

The amalgamated sample, which includes the main plus nonresponse study samples, with proper weighting, will provide unbiased estimates because the amalgamated sample will represent the entire population. Statistical tests that compare these unbiased estimates and estimates obtained solely from the (weighted) main sample will reveal whether the main sample estimates are significantly different from the unbiased estimates. We will refer to this method as the *amalgamated-main comparison.*

Nonresponse is of less concern if nonrespondents are not systematically different from the respondents in terms of the study variables. The second analysis focuses on this aspect using the super-population framework in which the two samples are assumed to be selected from hypothetical infinite populations of respondents and nonrespondents. This framework enables us to ignore the weights, simplifying the comparison. We performed *t* tests to determine whether the differences between estimates obtained from the unweighted data are significant. This method of comparison is termed the *unweighted comparison.*

The final set of analyses involved a series of logistic regressions in which participation status (main or initial respondents v. initial nonrespondents) was predicted using child age, disability category, and assessment scores. Significant coefficients for the assessment scores would provide evidence for potential bias due to nonresponse for those variables.

It should be noted that a significant difference in the unweighted analysis does not imply that the weighted main sample would be biased for the variable in question. It simply means that bias potential is greater. It is possible to eliminate the bias potential through effective nonresponse adjustment weighting. Therefore, greater emphasis should be given to the results of the amalgamated-main comparison.

Outcome Variables

Wave 1 demographic and direct assessment data were used to analyze nonresponse bias. Among the PEELS data, the direct assessment data are very key, as they will characterize the performance of preschoolers with disabilities and be used to model factors affecting that performance. Further, one might expect children's assessment performances to differ for districts that initially refused to participate in PEELS relative to those that initially accepted the PEELS invitation. Participating children completed a one-on-one assessment of school readiness with a trained assessor. The assessment included the following subtests:

- preLAS 2000 Simon Says, a measure of English/Spanish language ability;
- preLAS 2000 Art Show, a measure of English/Spanish language ability;
- Peabody Picture Vocabulary Test (PPVT), a measure of receptive language ability;
- Woodcock-Johnson III: Letter-Word Identification, a measure of pre-reading skill;

- Woodcock-Johnson III: Applied Problems, a measure of practical math skills;
- Woodcock-Johnson III: Quantitative Concepts-Concepts, a measure of conceptual math skills;
- Woodcock-Johnson III: Quantitative Concepts-Number Series;
- Leiter-R Attention Sustained Scale, a measure of attention;
- Individual Growth and Development Indicators (IGDI): Picture Naming, a measure of pre-reading skills;
- IGDI: Rhyming, a measure of pre-reading skills;
- IGDI: Alliteration, a measure of pre-reading skills;
- IGDI: Segment Blending, a measure of pre-reading skills; and
- Test of Early Math Skills, a measure of general math skills.

The above measures include a combination of performance (achievement) outcomes that we expect to be sensitive to the effects of programs and services that are provided to pre-elementary children and other variables (factors) that may help to explain performance. The PreLAS (Simon Says and Art Show) was used primarily to identify children needing a Spanish-language assessment rather than the Direct Assessment (in English). As such, these two measures were excluded from the nonresponse bias analysis. The PPVT, a measure of receptive language, is not considered to be an achievement measure. It was also excluded from the nonresponse bias analysis. Finally, the Test of Early Math Skills was thought to be largely duplicative of the several Woodcock-Johnson math measures already included in the analysis. Therefore, in order to reduce the complexity of the study, we elected to use only the Woodcock-Johnson measures. Thus, the remaining nine measures were used in the analysis.

Results

In the comparison of main and amalgamated sample estimates of child assessment scores, we assumed that the estimates obtained from the amalgamated sample were unbiased because they were based on the combination of main and nonresponse samples. To address the question of whether the main sample alone, which suffers a high rate of nonresponse, can produce unbiased estimates of the child assessment variables after weighting adjustment for nonresponses, we performed *t* tests on the differences of the estimates obtained from the amalgamated sample and the main sample. If a test result was significant for a variable, we interpreted the result as a piece of evidence to indicate a potential for bias in the main sample estimates for the variable. A non-significant result indicated a lack of such evidence. Tables C-2 through C-4 present the test results for nine outcome performance score variables[27] and eight additional demographic variables, including age, sex, and disability category.

In the following discussion, we will use a 5-percent significance level for all tests. The test results are given in terms of the *p*-value. If a *p*-value is greater than 5 percent, the test result (i.e., the comparison being examined), to which that *p*-value applies, is not statistically significant. Thus, for a comparison

[27] An Attention variable (Leiter-R) was constructed for each age group (3-, 4-, and 5-year-olds). The other eight variables were analyzed using age group as an independent variable.

yielding a *p*-value above 5 percent, the assumption is that there is no statistical difference between those means.

Comparisons Between the Weighted Main and Amalgamated Samples

First, we looked at the sex, age, and disability category distributions as presented in table C-2. The percentage of males in the amalgamated sample is 71.5 percent, which is slightly higher than the main sample estimate of 69.8 percent. The difference is not significant, with a 31.2 percent *p*-value. The percentage of each age group is also not significantly different between the two samples. The *p*-values range from 12.7 to 84.6 percent. No significant differences in individual disability categories were detected either.

Comparison of the two estimates of each score across the age groups is shown in table C-3. Among the 11 variables, only one variable, the WJLWSCORE (Letter-Word), had a significant difference, with a *p*-value of 3.2 percent. All other *p*-values were non-significant. In fact, most results were quite distant from the significance level of 5 percent, with the exception of the WJQCNSCORE (Quantitative Concepts: Number Series) variable, whose *p*-value (6.7 percent) was just over 5 percent.

When the data were analyzed by age group, no differences were significant. The ATTEN variables cannot be analyzed by age because they are already specific to a particular age. Results for these three variables are presented in table C-3. Results for the other assessment-by-age variables are presented in table C-4.

The *t* test results presented here, based on the amalgamated-main comparison, do not indicate any systematic bias in the main sample estimates. Even for the case of the WJLWSCORE (Letter-Word) variable where the overall age comparison yielded a statistically significant result, no significant difference was detected for the comparisons performed within age groups. This provides strong evidence that the main sample is unbiased for the great majority of the assessment variables considered in this study.

Comparisons Between the Unweighted Main and Nonresponse Samples

In the comparison of unweighted means from the main and nonresponse samples, one—WJAPSCORE—of the eight across-age comparisons revealed a significant difference. Among the 8 across-age comparisons and the 18 by-age comparisons, 3 of the by-age results yielded a significant difference—ATTEN4, WJLWSCORE age 4, and WJAPSCORE age 4. These results are provided in detail in tables C-5 and C-6.

While these results in isolation might raise some concerns about possible bias, particularly in cohort B (age 4), it is important to remember that the analyses were unweighted, and weighting is designed to reflect the sampling probability as well as reduce bias due to nonresponse.

Grouped Overall Comparisons

If we look at the results from the view point of overall comparisons, we can make even stronger statements about such comparisons than about individual comparisons. We performed chi-square tests to compare the overall distributions of age and disability. For the age distribution, the difference between the

amalgamated and main samples is strongly insignificant at a *p*-value of 79 percent. Similarly, the difference in the disability distribution in the two samples is insignificant with a *p*-value of 69 percent.

The Bonferroni inequality is often used to perform multiple comparisons. If we perform a family of *t* tests to compare *k* pairs of means with a significance level *a* for each of the *k* individual *t* tests, then the overall significance level (type I error) of the family of *t* tests is at most *ka*. For example, if $k = 10$ and the *ka* is set at 5 percent, then $a = 0.5$ percent.

If we apply this procedure to the result given in table C-3 with an overall significance level of 5 percent, we can say that the differences in the 11 pairs of means are collectively insignificant. We can say the same for the result presented in table C-4 even more forcefully. Furthermore, the Bonferroni procedure enables us to claim that unweighted comparisons shown in tables C-5 and C-6 are not significantly different either in terms of overall comparison.

Logistic Regression Results

Logistic regression analysis was used to examine whether participation status depends on the assessment scores. Dependency indicates possible bias in the score variables. Since the participation status variable is dichotomous, we can examine such dependency using logistic regression, where we use participation status as the dependent variable and assessment scores, disability category, and age as independent variables. By adding age and disability category in the regression models, the dependency is studied by subgroups of age and disability category.

Researchers tried to put as many score variables as possible together in a single model. However, since many score variables are age dependent, we had to limit the age groups permissible in each model. Furthermore, for some scores (e.g., IGDI Alliteration and Rhyming scores), although the tests shared a common age group, we could not estimate the regression coefficients when the tests were placed in a single model. This occurred because the score variables are defined not only based on age but also based on other differing restrictions and this, in turn, created many cases with missing values on one of the score variables. Separate models were developed for those variables. In every model, assessment scores were insignificant predictors of participation status (see tables C-7-A through 7-H).

Conclusions

Based on the three sets of analyses presented here, we conclude that there is little evidence of response bias in the PEELS main sample data. While a few individual comparisons of unweighted data were significantly different, the comparisons of the weighted data were not, in particular when run by age. Furthermore, even those significantly different individual comparisons were not significant as a collective group. This suggests that the weights have eliminated bias in the unweighted main sample. In addition, none of the regressions indicated that assessment scores were significant predictors of participation status. Based on this evidence, we believe no systematic differences exist between the main and nonresponse bias study samples.

Table C-2. Main and amalgamated sample comparison of sex, age, and disability categories

Variable Name	Main		Amalgamated		Difference on main and amalgamated sample est					
	N	est	*N*	est	est	*s.e.*	Lower C.L.	Upper C.L.	*t* test *p*-value	Significant?
SEX_1	2,242	0.698	2,426	0.715	-0.018	0.017	-0.052	0.017	0.312	No
SEX_2	2,242	0.302	2,426	0.285	0.018	0.017	-0.017	0.052	0.312	No
AGE_3	2,242	0.182	2,426	0.194	-0.012	0.008	-0.027	0.003	0.127	No
AGE_4	2,242	0.368	2,426	0.358	0.010	0.013	-0.017	0.036	0.471	No
AGE_5	2,242	0.418	2,426	0.421	-0.003	0.013	-0.028	0.023	0.846	No
DDCAT_1	2,242	0.345	2,426	0.331	0.014	0.032	-0.050	0.077	0.666	No
DDCAT_2	2,242	0.505	2,426	0.491	0.014	0.028	-0.042	0.070	0.622	No
DDCAT_3	2,242	0.030	2,426	0.026	0.004	0.009	-0.014	0.021	0.690	No
DDCAT_4	2,242	0.035	2,426	0.051	-0.016	0.013	-0.042	0.010	0.229	No
DDCAT_5	2,242	0.046	2,426	0.059	-0.012	0.015	-0.043	0.018	0.426	No
DDCAT_6	2,242	0.006	2,426	0.006	0.001	0.003	-0.005	0.006	0.873	No
DDCAT_7	2,242	0.033	2,426	0.037	-0.004	0.010	-0.023	0.016	0.704	No

Table C-3. Main and amalgamated sample comparison of the means of child assessment scores

	Main		Amalgamated		Difference					
Variable Name	***N***	**est**	***N***	**est**	**est**	***s.e.***	**Lower C.L.**	**Upper C.L.**	***t* test *p*-value**	**Significant?**
WJQCCScore	807	7.37	863	7.30	0.06	0.28	-0.49	0.62	0.822	No
WJQCNSScore	807	3.55	863	3.16	0.40	0.22	-0.03	0.82	0.067	No
WJAPScore	2,242	10.38	2,426	10.10	0.29	0.24	-0.18	0.76	0.225	No
WJLWScore	2,239	7.93	2,423	7.50	0.43	0.20	0.04	0.82	0.032	No
IGDIPNScore	2,014	14.70	2,178	15.04	-0.34	0.32	-0.98	0.30	0.296	No
IGDIAScore	720	4.96	775	5.07	-0.11	0.34	-0.77	0.56	0.751	No
IGDIRScore	774	6.55	823	6.67	-0.12	0.49	-1.08	0.84	0.812	No
IGDISBScore	1,562	10.17	1,681	10.69	-0.52	0.52	-1.56	0.51	0.317	No
ATTEN3	533	9.15	586	8.96	0.18	0.31	-0.44	0.81	0.557	No
ATTEN4	859	9.07	930	8.70	0.37	0.25	-0.12	0.86	0.139	No
ATTEN5	776	9.30	826	9.59	-0.29	0.38	-1.05	0.47	0.445	No

Table C-4. Main and amalgamated sample comparison of the means of child assessment scores, by age group

		Main		Amalgamated		Difference					
Variable Name	Age group	*N*	est	*N*	est	est	*s.e.*	Lower C.L.	Upper C.L.	*t* test *p*-value	Significant?
WJAPScore	Age 3	587	5.19	641	5.17	0.01	0.43	-0.83	0.86	0.973	No
	Age 4	848	9.11	922	8.68	0.43	0.41	-0.39	1.24	0.302	No
	Age 5	749	13.28	801	13.19	0.09	0.43	-0.75	0.94	0.825	No
WJLWScore	Age 3	586	4.10	640	4.24	-0.14	0.45	-1.03	0.75	0.756	No
	Age 4	846	5.98	920	5.56	0.42	0.27	-0.12	0.97	0.124	No
	Age 5	749	10.84	801	10.22	0.62	0.42	-0.21	1.45	0.142	No
IGDIPNScore	Age 3	477	10.95	519	11.56	-0.61	0.46	-1.51	0.29	0.183	No
	Age 4	773	13.81	842	13.41	0.40	0.51	-0.60	1.41	0.429	No
	Age 5	711	16.50	760	17.45	-0.94	0.59	-2.10	0.22	0.110	No
IGDIAScore	Age 4	254	3.48	279	3.26	0.22	0.32	-0.40	0.85	0.486	No
	Age 5	426	5.48	454	5.93	-0.45	0.62	-1.66	0.77	0.470	No
IGDIRScore	Age 4	302	5.11	320	4.97	0.14	0.27	-0.38	0.67	0.596	No
	Age 5	431	7.02	459	7.31	-0.30	0.73	-1.73	1.14	0.683	No
IGDISBScore	Age 4	785	7.30	852	7.60	-0.30	0.54	-1.37	0.77	0.579	No
	Age 5	719	12.06	768	12.61	-0.55	0.90	-2.32	1.23	0.545	No

Table C-5. Main and nonresponse sample comparison of the unweighted means of child assessment scores

	Main		Nonresponse		Difference					
Variable Name	***N***	**est**	***N***	**est**	**est**	***s.e.***	**Lower C.L.**	**Upper C.L.**	***t* test *p*-value**	**Significant?**
M_WJQCCScore	807	7.24	56	7.16	0.08	0.450	-0.80	0.96	0.843	No
M_WJQCNSScore	807	3.34	56	2.91	0.43	0.413	-0.38	1.24	0.293	No
M_WJAPScore	2,242	9.68	184	8.50	1.18	0.457	0.29	2.08	0.010	No
M_WJLWScore	2,239	7.10	184	6.29	0.81	0.441	-0.06	1.67	0.064	No
M_IGDIPNScore	2,014	14.50	164	14.61	-0.11	0.509	-1.11	0.89	0.836	No
M_IGDIAScore	720	4.89	55	4.60	0.29	0.559	-0.81	1.39	0.556	No
M_IGDIRScore	774	6.42	49	6.35	0.07	0.680	-1.26	1.40	0.919	No
M_IGDISBScore	1,562	9.91	119	9.90	0.01	0.830	-1.62	1.64	0.989	No
M_ATTEN3	533	9.18	53	8.58	0.59	0.463	-0.32	1.50	0.283	No
M_ATTEN4	859	9.26	71	8.21	1.05	0.439	0.19	1.91	0.009	No
M_ATTEN5	776	9.50	53	9.40	0.10	0.561	-1.00	1.20	0.868	No

Table C-6. Main and nonresponse sample comparison of the unweighted means of child assessment scores, by age

		Main		Nonresponse		Difference					
Variable Name	Age group	*N*	est	*N*	est	est	*s.e.*	Lower C.L.	Upper C.L.	*t* test *p*-value	Significant?
M_WJAPScore	Age 3	587	5.16	54	5.17	-0.01	0.615	-1.21	1.20	0.992	No
	Age 4	848	9.31	74	7.65	1.66	0.610	0.47	2.86	0.009	No
	Age 5	749	13.14	52	12.83	0.31	0.780	-1.22	1.84	0.698	No
M-WJLWScore	Age 3	586	4.03	54	4.04	-0.01	0.539	-1.06	1.05	0.994	No
	Age 4	846	5.99	74	4.96	1.03	0.542	-0.04	2.09	0.035	No
	Age 5	749	10.20	52	10.12	0.08	0.900	-1.68	1.86	0.928	No
M_IGDIPNScore	Age 3	477	10.93	42	11.71	-0.78	0.869	-2.49	0.92	0.324	No
	Age 4	773	14.24	69	13.42	0.82	0.733	-0.62	2.26	0.282	No
	Age 5	711	16.82	49	18.43	-1.61	0.888	-3.35	0.14	0.069	No
M_IGDIAScore	Age 4	254	3.70	25	3.20	0.50	0.621	-0.72	1.72	0.289	No
	Age 5	426	5.41	28	5.75	-0.34	0.847	-2.00	1.32	0.676	No
M_IGDIRScore	Age 4	302	5.13	18	4.67	0.46	0.963	-1.43	2.36	0.587	No
	Age 5	431	7.05	28	7.43	-0.38	0.924	-2.19	1.44	0.706	No
M_IGDISBScore	Age 4	785	7.43	67	7.28	0.15	0.887	-1.59	1.89	0.850	No
	Age 5	719	12.06	49	12.78	-0.72	1.388	-3.44	2.01	0.617	No

Table C-7-A. Logistic regression results for model of Woodcock-Johnson III: Quantitative Concepts scores

HYPOTHESIS TESTING RESULTS: 863 (UNWEIGHTED)

TEST	F VALUE	NUM. DF	DENOM. DF	PROB>F	NOTE
OVERALL FIT	0.413	8	114	0.911	
WJQCCScore	1.914	1	121	0.169	
WJQCNSScore	2.436	1	121	0.121	
ddiscat2[7]	0.186	6	116	0.98	

ESTIMATED FULL SAMPLE REGRESSION COEFFICIENTS

PARAMETER	PARAMETER ESTIMATE	STANDARD ERROR OF ESTIMATE	TEST FOR H0: PARAMETER=0	PROB>|T|	COMMENT
INTERCEPT	0.3	1.279	0.237	0.813	
WJQCCScore	-0.11	0.078	-1.384	0.169	
WJQCNSScore	0.13	0.082	1.561	0.121	
ddiscat2.1	-0.13	0.804	-0.158	0.874	
ddiscat2.2	0.06	0.922	0.06	0.952	
ddiscat2.3	0.55	34.731	0.016	0.987	Unstable Standard Error
ddiscat2.4	-0.5	1.351	-0.372	0.711	
ddiscat2.5	0.32	2.068	0.156	0.877	
ddiscat2.6	0.32	32.915	0.01	0.992	Unstable Standard Error

Table C-7-B. Logistic regression results for model of Woodcock Johnson III Letter-Word and Applied Problems scores

HYPOTHESIS TESTING RESULTS: 2178 (UNWEIGHTED)

TEST	F VALUE	NUM. DF	DENOM. DF	PROB>F
OVERALL FIT	2.1327	11	111	0.0234
ddiscat2[7]	0.5529	6	116	0.7669
WJLWScore	2.6736	1	121	0.1046
WJAPScore	0.5406	1	121	0.4636
IGDIPNScore	1.4604	1	121	0.2292
CHLDAGE2[3]	0.5636	2	120	0.5707

ESTIMATES FULL REGRESSION COEFFICIENTS

PARAMETER	PARAMETER ESTIMATE	STANDARD ERROR OF ESTIMATE	TEST FOR H0: PARAMETER=0	PROB>|T|
INTERCEPT	-0.18	1.1105	-0.1638	0.8702
ddiscat2.1	0.16	0.6333	0.2587	0.7963
ddiscat2.2	0.29	0.6419	0.4593	0.6469
ddiscat2.3	-0.13	1.2519	-0.1015	0.9193
ddiscat2.4	-0.73	1.1091	-0.6582	0.5117
ddiscat2.5	-0.27	1	-0.2701	0.7875
ddiscat2.6	0.81	32.9739	0.0245	0.9805
WJLWScore	0.03	0.0208	1.6351	0.1046
WJAPScore	0.03	0.0361	0.7353	0.4636
IGDIPNScore	-0.05	0.0384	-1.2085	0.2292
CHLDAGE2.1	0.14	0.7784	0.1809	0.8568
CHLDAGE2.2	0.35	0.5473	0.635	0.5266

Table C-7-C. Logistic regression results for model of IGDI Alliteration scores

HYPOTHESIS TESTING RESULTS: 775 (UNWEIGHTED)

TEST	F VALUE	NUM. DF	DENOM. DF	PROB>F
OVERALL FIT	0.043	5	117	0.999
ddiscat3[4]	0.013	3	119	0.998
CHLDAGE2[2]	0.045	1	121	0.832
IGDIAScore	0.216	1	121	0.643

ESTIMATED FULL SAMPLE REGRESSION COEFFICIENTS

PARAMETER	PARAMETER ESTIMATE	STANDARD ERROR OF ESTIMATE	TEST FOR H0: PARAMETER=0	PROB>|T|
INTERCEPT	0.25	1.955	0.126	0.9
ddiscat3.1	-0.17	1.831	-0.095	0.924
ddiscat3.2	-0.1	1.901	-0.054	0.957
ddiscat3.3	-0.14	2.352	-0.058	0.954
CHLDAGE2.1	-0.14	0.64	-0.213	0.832
IGDIAScore	-0.03	0.07	-0.465	0.643

Table C-7-D. Logistic regression results for model of IGDI Rhyming scores

HYPOTHESIS TESTING RESULTS: 823 (UNWEIGHTED)

TEST	F VALUE	NUM. DF	DENOM. DF	PROB>F	NOTE
OVERALL FIT	0.304	5	117	0.91	
ddiscat3[4]	0.201	3	119	0.896	
CHLDAGE2[2]	0.157	1	121	0.693	
IGDIRScore	0.195	1	121	0.66	

ESTIMATED FULL SAMPLE REGRESSION COEFFICIENTS

PARAMETER	PARAMETER ESTIMATE	STANDARD ERROR OF ESTIMATE	TEST FOR H0: PARAMETER=0	PROB>|T|	COMMENT
INTERCEPT	0.59	1.47	0.399	0.691	
ddiscat3.1	-0.11	1.728	-0.066	0.948	
ddiscat3.2	-0.5	1.538	-0.325	0.746	
ddiscat3.3	-0.55	34.21	-0.016	0.987	Unstable Standard Error
CHLDAGE2.1	0.28	0.697	0.396	0.693	
IGDIRScore	-0.03	0.067	-0.442	0.66	

Table C-7-E. Logistic regression results for model of IGDI Segment Blending scores

HYPOTHESIS TESTING RESULTS: 1681 (UNWEIGHTED)

TEST	F VALUE	NUM. DF	DENOM. DF	PROB>F
OVERALL FIT	0.639	5	117	0.67
CHLDAGE2[2]	0.076	1	121	0.783
ddiscat3[4]	0.229	3	119	0.876
IGDISBScore	0.441	1	121	0.508

ESTIMATED FULL SAMPLE REGRESSION COEFFICIENTS

PARAMETER	PARAMETER ESTIMATE	STANDARD ERROR OF ESTIMATE	TEST FOR H0: PARAMETER=0	PROB>|T|
INTERCEPT	-0.25	0.794	-0.315	0.753
CHLDAGE2.1	0.15	0.555	0.276	0.783
ddiscat3.1	0.28	0.873	0.32	0.749
ddiscat3.2	0.41	0.771	0.538	0.591
ddiscat3.3	1.28	1.716	0.746	0.457
IGDISBScore	-0.01	0.022	-0.664	0.508

Table C-7-F. Logistic regression results for model of Leiter-R Attention Sustained scores, age 3

HYPOTHESIS TESTING RESULTS: 586 (UNWEIGHTED)

TEST	F VALUE	NUM. DF	DENOM. DF	PROB>F
OVERALL				
FIT	0.631	4	118	0.641
ddiscat3[4]	0.515	3	119	0.672
ATTEN3	0.618	1	121	0.433

ESTIMATED FULL SAMPLE REGRESSION COEFFICIENTS

PARAMETER	PARAMETER ESTIMATE	STANDARD ERROR OF ESTIMATE	TEST FOR H0: PARAMETER=0	PROB>|T|
INTERCEPT	-1.58	1.727	-0.915	0.362
ddiscat3.1	0.66	1.35	0.486	0.628
ddiscat3.2	1.19	1.513	0.785	0.434
ddiscat3.3	-0.37	2.354	-0.156	0.876
ATTEN3	0.06	0.073	0.786	0.433

Table C-7-G. Logistic regression results for model of Leiter-R Attention Sustained scores, age 4

HYPOTHESIS TESTING RESULTS: 929 (UNWEIGHTED)

TEST	F VALUE	NUM. DF	DENOM. DF	PROB>F
OVERALL				
FIT	1.005	4	118	0.408
ddiscat3[4]	0.426	3	119	0.734
ATTEN4	3.082	1	121	0.082

ESTIMATED FULL SAMPLE REGRESSION COEFFICIENTS

PARAMETER	PARAMETER ESTIMATE	STANDARD ERROR OF ESTIMATE	TEST FOR H0: PARAMETER=0	PROB>|T|
INTERCEPT	-1.59	1.6	-0.991	0.324
ddiscat3.1	0.67	1.476	0.452	0.652
ddiscat3.2	1.1	1.477	0.746	0.457
ddiscat3.3	1.64	1.828	0.898	0.371
ATTEN4	0.1	0.059	1.756	0.082

Table C-7-H. Logistic regression results for model of Leiter-R Attention Sustained scores, age 5

HYPOTHESIS TESTING RESULTS: 829 (UNWEIGHTED)

TEST	F VALUE	NUM. DF	DENOM. DF	PROB>F	NOTE
OVERALL					
FIT	0.139	4	118	0.967	
ddiscat3[4]	0.032	3	119	0.992	
ATTEN5	0.459	1	121	0.5	

ESTIMATED FULL SAMPLE REGRESSION COEFFICIENTS

PARAMETER	PARAMETER ESTIMATE	STANDARD ERROR OF ESTIMATE	TEST FOR H0: PARAMETER=0	PROB>|T|	COMMENT
INTERCEPT	0.19	1.104	0.176	0.861	
ddiscat3.1	0.16	0.971	0.169	0.866	
ddiscat3.2	0.27	1.022	0.261	0.795	
ddiscat3.3	0.57	34.718	0.016	0.987	Unstable Standard Error
ATTEN5	-0.04	0.065	-0.677	0.5	

Appendix D: Standard Error Tables

The tables in Appendix D contain standard errors for the corresponding tables in the main body of the report. For example, table D-12 contains the standard errors for table 12.

Table D-12. Standard errors for the percentage of young children who received preschool special education services and had or did not have IEPs in 2003-04 and 2004-05

		2004-05		
		Total	IEP/IFSP	No IEP/IFSP
2003-04	Total		1.1	1.1
	IEP/IFSP	0.7	1.1	1.0
	No IEP/IFSP	0.7	0.4	0.7

SOURCE: U.S. Department of Education, National Center for Special Education Research, Pre-Elementary Education Longitudinal Study (PEELS), "Elementary School Teacher Questionnaire," "Kindergarten Teacher Questionnaire," "Early Childhood Teacher Questionnaire," and "Parent Interview."

Table D-13. Standard errors for the percentage of young children who received preschool special education services during the 2003-04 school year and were declassified in 2004-05

	Declassified
Expected percentage based on the total sample	1.0
Gender	1.2
Male	2.1
Female	1.0
Ethnicity	
Black	1.8
Hispanic	2.3
White	1.3
Family income	
$20,000 or less	2.8
$20,001-$40,000	1.9
More than $40,000	1.6
Metropolitan status	
Urban	2.1
Suburban	1.3
Rural	2.1
District size	
Very large	2.2
Large	2.2
Medium	2.2
Small	1.6
District wealth	
High	1.7
Medium	1.7
Low	2.7
Very low	2.2

SOURCE: U.S. Department of Education, National Center for Special Education Research, Pre-Elementary Education Longitudinal Study (PEELS), "Elementary School Teacher Questionnaire," "Kindergarten Teacher Questionnaire," "Early Childhood Teacher Questionnaire," and "Parent Interview."

Table D-14. Standard errors for the percentage of young children who received preschool special education services during the 2003-04 school year and were declassified in 2004-05, by disability

	Disability									
	Expected percentage based on the total sample	AU	DD	ED	LD	MR	OI	OHI	SLI	LI
Percentage of children in each disability group who were declassified	1.0	0.4	3.7	0.9	0.6	‡	0.9	‡	5.3	2.1

‡ Reporting standards not met.
NOTE: AU = Autism; DD = Developmental delay; ED = Emotional disturbance; LD = Learning disability; MR = Mental retardation; OI = Orthopedic impairment; OHI = Other health impairment; SLI = Speech or language impairment; LI = Low incidence.
SOURCE: U.S. Department of Education, National Center for Special Education Research, Pre-Elementary Education Longitudinal Study (PEELS), "Elementary School Teacher Questionnaire," "Kindergarten Teacher Questionnaire," "Early Childhood Teacher Questionnaire," and "Parent Interview."

Table D-15. Standard errors for the percentage of young children who received preschool special education services during the 2003-04 school year and were declassified in 2004-05, by transition status

	Transition status					
	Expected percentage based on the total sample	Remained in preschool	Transitioned from preschool to kindergarten	Transitioned from kindergarten to first grade	Remained out of school	Other status
Percentage of children in each transition status who were declassified	1.0	1.1	1.4	4.3	8.7	7.2

SOURCE: U.S. Department of Education, National Center for Special Education Research, Pre-Elementary Education Longitudinal Study (PEELS), "Elementary School Teacher Questionnaire," "Kindergarten Teacher Questionnaire," "Early Childhood Teacher Questionnaire," and "Parent Interview."

Table D-16 Standard errors for the mean performance of young children who received preschool special education services during the 2003-04 school year on tests of emerging literacy and early math skills, by eligibility status

	Letter-Word Identification	Applied Problems	PPVT
Total	0.5	0.7	0.6
Remained eligible	0.6	0.7	0.6
Declassified	1.0	1.2	1.3

SOURCE: U.S. Department of Education, National Center for Special Education Research, Pre-Elementary Education Longitudinal Study (PEELS), "Woodcock-Johnson III: Letter-Word Identification," "Woodcock-Johnson III: Applied Problems," "Peabody Picture Vocabulary Test III–R."

Table D-17. Standard errors for the mean performance of young children who received preschool special education services during the 2003-04 and 2004-05 school years on tests of emerging literacy and early math skills, by reclassification status

	Letter-Word Identification	Applied Problems	PPVT
Total	0.7	1.0	0.7
Reclassified	1.5	1.7	1.4
Not reclassified	0.8	1.1	0.7

SOURCE: U.S. Department of Education, National Center for Special Education Research, Pre-Elementary Education Longitudinal Study (PEELS), "Woodcock-Johnson III: Letter-Word Identification," "Woodcock-Johnson III: Applied Problems," "Peabody Picture Vocabulary Test III–R."

Table D-18. Standard errors for the percentage of young children who received preschool special education services whose disability classification remained the same from 2003-04 to 2004-05

Disability classification	
Total	1.2
Autism	2.9
Developmental delay	3.2
Emotional disturbance	9.2
Learning disability	6.8
Mental retardation	5.8
Orthopedic impairment	11.9
Other health impairment	10.2
Speech or language impairment	1.3
Low-incidence disability	5.5

NOTE: Percentages do not include children who were declassified between 2003-04 and 2004-05.
SOURCE: U.S. Department of Education, National Center for Special Education Research, Pre-Elementary Education Longitudinal Study (PEELS), "Early Childhood Teacher Questionnaire," "Kindergarten Teacher Questionnaire," "Elementary School Teacher Questionnaire," "Parent Interview" previously unpublished tabulation (August 2006).

Table D-19. Standard errors for the percentage of young children who received preschool special education services who received specific services through their school system, by school year: School years 2003-04 and 2004-05

	2003-04	2004-05
Adaptive physical education	1.1	1.2
Assistive technology services/devices	1.1	1.3
Audiology	1.4	1.2
Augmentative or alternative communication system	1.0	0.8
Behavior management program	1.5	0.9
Learning strategies/study skills assistance	2.0	1.0
Occupational therapy	1.6	1.5
One-to-one paraeducators/assistant	0.8	1.1
Physical therapy	1.2	1.3
Service coordination/case management	2.4	1.0
Social work services	1.0	0.9
Special transportation because of disability	1.4	1.2
Specialized computer software or hardware	1.1	0.6
Speech or language therapy	1.5	1.6
Training, counseling, or other supports/services for family	1.5	0.6
Tutoring/remediation by a special education teacher	1.6	1.1
Other services	1.3	1.0

NOTES: Other services include health services; instruction in American Sign Language, Manual English, Cued Speech, or Braille; mental health services; reader or interpreter; vision services; and other services specified by the respondent. Denominators do not include children who were declassified from special education, so percentages include only children with an IEP.

SOURCE: U.S. Department of Education, National Center for Special Education Research, Pre-Elementary Education Longitudinal Study (PEELS), "Elementary School Teacher Questionnaire," "Kindergarten Teacher Questionnaire," and "Early Childhood Teacher Questionnaire," previously unpublished tabulation (April 2006).

Table D-20. Standard errors for the mean number of services provided to young children who received preschool special education services, by age cohort, gender, and disability: School years 2003-04 and 2004-05

	2003-04	2004-05
Total	0.1	0.1
Age cohort		
Cohort A	0.1	0.
Cohort B	0.1	0.1
Cohort C	0.1	0.1
Gender		
Male	0.1	0.1
Female	0.2	0.2
Disability		
Autism	0.4	0.3
Developmental delay	0.1	0.1
Emotional disturbance	0.5	0.4
Learning disability	0.3	0.3
Mental retardation	0.4	0.3
Orthopedic impairment	0.8	0.5
Other health impairment	0.7	0.5
Speech or language impairment	0.1	0.1
Low-incidence disability	0.6	0.4

SOURCE: U.S. Department of Education, National Center for Special Education Research, Pre-Elementary Education Longitudinal Study (PEELS), "Elementary School Teacher Questionnaire," "Kindergarten Teacher Questionnaire," and "Early Childhood Teacher Questionnaire," previously unpublished tabulation (April 2006).

Table D-21. Standard error for the mean hours per week that young children who received preschool special education services spent in various educational settings, by age cohort and school year: School years 2003-04 and 2004-05

	2003-04	2004-05
Total		
Regular education classroom	0.5	0.5
Special education setting	0.5	0.4
Therapy setting	0.1	#
Non-special education setting outside the classroom for remedial or special assistance	0.2	#
Home instruction	#	#
Cohort A		
Regular education classroom	0.4	0.4
Special education setting	0.5	0.6
Therapy setting	0.1	#
Non-special education setting outside the classroom for remedial or special assistance	#	#
Home instruction	0.1	#
Cohort B		
Regular education classroom	0.6	0.5
Special education setting	0.6	0.5
Therapy setting	0.1	0.1
Non-special education setting outside the classroom for remedial or special assistance	#	#
Home instruction	0.1	0.1
Cohort C		
Regular education classroom	1.0	0.8
Special education setting	0.8	0.4
Therapy setting	0.2	0.1
Non-special education setting outside the classroom for remedial or special assistance	0.4	0.1
Home instruction	0.1	0.1

Rounds to zero

SOURCE: U.S. Department of Education, National Center for Special Education Research, Pre-Elementary Education Longitudinal Study (PEELS), "Elementary School Teacher Questionnaire," "Kindergarten Teacher Questionnaire," and "Early Childhood Teacher Questionnaire," previously unpublished tabulation (April 2006).

Table D-22. Standard error for the percentage of young children who received preschool special education services whose parents were satisfied with special education services to various degrees: School years 2003-2004 and 2004-05

	2003-04	2004-05
Very satisfied	1.9	2.1
Satisfied	1.7	1.7
Dissatisfied	0.7	0.8
Very dissatisfied	0.3	0.3

SOURCE: U.S. Department of Education, National Center for Special Education Research, Pre-Elementary Education Longitudinal Study (PEELS), "Parent Interview," previously unpublished tabulation (April 2006).

Table D-23. Standard errors for the mean performance of young children who received preschool special education services on Woodcock-Johnson III: Letter-Word Identification, by age cohort: School years 2003–04 and 2004-05

	Total	Cohort A	Cohort B	Cohort C
2003-04	0.5	1.1	0.8	0.6
2004-05	0.5	0.8	0.9	0.7

SOURCE: U.S. Department of Education, National Center for Special Education Research, Pre-Elementary Education Longitudinal Study (PEELS), "Woodcock-Johnson III: Letter-Word Identification."

Table D-24. Standard errors for the mean performance of young children who received preschool special education services on Woodcock-Johnson III: Letter-Word Identification, by Wave 1 primary disability: School years 2003-04 and 2004-05

	AU	DD	ED	LD	MR	OI	OHI	SLI	LI
Total									
2003-04	5.1	0.9	3.1	3.1	5.0	2.3	3.8	0.7	1.7
2004-05	5.4	1.0	3.0	2.8	5.6	2.1	4.9	0.6	2.7

NOTE: AU = Autism; DD = Developmental delay; ED = Emotional disturbance; LD = Learning disability; MR = Mental retardation; OI = Orthopedic impairment; OHI = Other health impairment; SLI = Speech or language impairment; LI = Low incidence.
SOURCE: U.S. Department of Education, National Center for Special Education Research, Pre-Elementary Education Longitudinal Study (PEELS), "Woodcock-Johnson III: Letter-Word Identification."

Table D-25. Standard errors for the mean performance of young children who received preschool special education services on Peabody Picture Vocabulary Test III, by age cohort: School years 2003-04 and 2004-05

	Total	Cohort A	Cohort B	Cohort C
2003-04	0.7	0.7	0.7	1.3
2004-05	0.6	0.6	0.8	1.0

SOURCE: U.S. Department of Education, National Center for Special Education Research, Pre-Elementary Education Longitudinal Study (PEELS), "Peabody Picture Vocabulary Test III."

Table D-26. Standard errors for the mean performance of young children who received preschool special education services on Peabody Picture Vocabulary Test III, by Wave 1 primary disability: School years 2003-04 and 2004-05

	AU	DD	ED	LD	MR	OI	OHI	SLI	LI
Total									
2003-04	3.0	1.1	2.6	1.4	4.0	2.6	3.7	0.7	1.8
2004-05	3.8	0.7	1.5	1.9	5.2	3.6	3.0	0.7	1.7

NOTE: AU = Autism; DD = Developmental delay; ED = Emotional disturbance; LD = Learning disability; MR = Mental retardation; OI = Orthopedic impairment; OHI = Other health impairment; SLI = Speech or language impairment; LI = Low incidence.
SOURCE: U.S. Department of Education, National Center for Special Education Research, Pre-Elementary Education Longitudinal Study (PEELS), "Peabody Picture Vocabulary Test III."

Table D-27. Standard errors for the mean performance of young children who received preschool special education services on Woodcock-Johnson III: Applied Problems, by age cohort: School years 2003-04 and 2004-05

	Total	Cohort A	Cohort B	Cohort C
2003-04	0.7	1.1	1.1	1.1
2004-05	0.7	0.9	1.2	1.1

SOURCE: U.S. Department of Education, National Center for Special Education Research, Pre-Elementary Education Longitudinal Study (PEELS), "Woodcock-Johnson III: Applied Problems."

Table D-28. Standard errors for the mean performance of young children who received preschool special education services on Woodcock-Johnson III: Applied Problems, by gender: School years 2003-04 and 2004-05

	Male	Female
Total		
2003-04	0.8	1.1
2004-05	0.7	1.2

SOURCE: U.S. Department of Education, National Center for Special Education Research, Pre-Elementary Education Longitudinal Study (PEELS), "Woodcock-Johnson III: Applied Problems."

Table D-29. Standard errors for the mean performance of young children who received preschool special education services on Woodcock-Johnson III: Applied Problems, by Wave 1 primary disability: School years 2003-04 and 2004-05

	AU	DD	ED	LD	MR	OI	OHI	SLI	LI
Total									
2003-04	4.2	1.2	3.3	2.0	3.6	2.8	5.2	0.8	3.0
2004-05	5.3	0.9	3.5	2.5	4.5	2.7	5.7	0.6	4.1

NOTE: AU = Autism; DD = Developmental delay; ED = Emotional disturbance; LD = Learning disability; MR = Mental retardation; OI = Orthopedic impairment; OHI = Other health impairment; SLI = Speech or language impairment; LI = Low incidence.
SOURCE: U.S. Department of Education, National Center for Special Education Research, Pre-Elementary Education Longitudinal Study (PEELS), "Woodcock-Johnson III: Applied Problems."

Table D-30. Standard errors for the mean performance of young children who received preschool special education services on Woodcock-Johnson III: Quantitative Concepts, by age cohort: School years 2003-04 and 2004-05

	Total	Cohort A	Cohort B	Cohort C
2003-04	0.8	†	†	0.8
2004-05	0.7	†	1.0	0.9

† Not applicable.
SOURCE: U.S. Department of Education, National Center for Special Education Research, Pre-Elementary Education Longitudinal Study (PEELS), "Woodcock-Johnson III: Quantitative Concepts."

Table D-31. Standard errors for mean teacher ratings of young children who received preschool special education services on the Social Skills subscale of the Preschool and Kindergarten Behavior Scale, by age cohort: School years 2003-04 and 2004-05

	Total	Cohort A	Cohort B	Cohort C
2003-04	0.5	0.9	1.0	1.2
2004-05	0.6	0.9	0.7	1.1

SOURCE: U.S. Department of Education, National Center for Special Education Research, Pre-Elementary Education Longitudinal Study (PEELS), "Social Skills subscale of the Preschool and Kindergarten Behavior Scale."

Table D-32. Standard errors for mean teacher ratings of young children who received preschool special education services on the Social Skills subscale of the Preschool and Kindergarten Behavior Scale, by gender: School years 2003-04 and 2004-05

	Male	Female
Total		
2003-04	0.6	1.6
2004-05	0.6	1.2

SOURCE: U.S. Department of Education, National Center for Special Education Research, Pre-Elementary Education Longitudinal Study (PEELS), "Social Skills subscale of the Preschool and Kindergarten Behavior Scale."

Table D-33. Standard errors for mean teacher ratings of young children who received preschool special education services on the Social Skills subscale of the Preschool and Kindergarten Behavior Scale, by Wave 1 primary disability: School years 2003-04 and 2004-05

	AU	DD	ED	LD	MR	OI	OHI	SLI	LI
Total									
2003-04	3.2	1.2	5.1	4.2	4.4	3.1	5.6	0.7	4.5
2004-05	2.4	1.1	4.5	4.7	4.3	3.4	5.0	0.7	4.7

NOTE: AU = Autism; DD = Developmental delay; ED = Emotional disturbance; LD = Learning disability; MR = Mental retardation; OI = Orthopedic impairment; OHI = Other health impairment; SLI = Speech or language impairment; LI = Low incidence.
SOURCE: U.S. Department of Education, National Center for Special Education Research, Pre-Elementary Education Longitudinal Study (PEELS), "Social Skills subscale of the Preschool and Kindergarten Behavior Scale."

Table D-34. Standard errors for mean teacher ratings of young children who received preschool special education services on the Vineland Adaptive Behavior Scales, Motor Skills Domain, by age cohort: School years 2003-04 and 2004-05

	Total	Cohort A	Cohort B	Cohort C
2003-04	0.9	0.9	1.1	1.6
2004-05	1.0	1.0	1.0	1.6

SOURCE: U.S. Department of Education, National Center for Special Education Research, Pre-Elementary Education Longitudinal Study (PEELS), "Vineland Adaptive Behavior Scales, Motor Skills Domain."

Table D-35. Standard errors for mean teacher ratings of young children who received preschool special education services and participated in an alternate assessment: ABAS II—Skill area scores (Teacher/Daycare Provider Form): School years 2003-04 and 2004-05

	Total
Communication	
2003-04	0.2
2004-05	0.1
Functional (pre) academics	
2003-04	0.1
2004-05	0.3
Health and safety	
2003-04	0.1
2004-05	0.2
Leisure	
2003-04	0.1
2004-05	0.2
Motor	
2003-04	0.2
2004-05	0.4
School living	
2003-04	0.2
2004-05	0.3
Self-care	
2003-04	0.2
2004-05	0.3
Self-direction	
2003-04	0.2
2004-05	0.3
Social	
2003-04	0.1
2004-05	0.2

SOURCE: U.S. Department of Education, National Center for Special Education Research, Pre-Elementary Education Longitudinal Study (PEELS), "ABAS-II—Skill area scores."

Table D-36. Standard errors for mean teacher ratings of young children who received preschool special education services and participated in an alternate assessment: ABAS II—Skill area scores (Teacher Form): School years 2003-04 and 2004-05

	Total
Communication	
2003-04	0.2
2004-05	0.1
Community use	
2003-04	0.3
2004-05	0.3
Functional academics	
2003-04	0.6
2004-05	0.3
Health and safety	
2003-04	0.5
2004-05	0.2
Leisure	
2003-04	0.3
2004-05	0.3
School living	
2003-04	0.6
2004-05	0.4
Self-care	
2003-04	0.7
2004-05	0.3
Self-direction	
2003-04	0.1
2004-05	0.1
Social	
2003-04	0.3
2004-05	0.2

SOURCE: U.S. Department of Education, National Center for Special Education Research, Pre-Elementary Education Longitudinal Study (PEELS), "ABAS-II—Skill area scores."

Appendix E: Number of Children Who Had Test Accommodations

Table E-1. Unweighted number of children who had various test accommodations in the PEELS Wave 2 direct assessment, by gender: School year 2004-05

	Male	Female
Abacus	‡	‡
Adaptive furniture	8	4
Communication device	‡	‡
Enlarged print	‡	‡
Familiar person administered test	‡	‡
Familiar person present	62	20
Multiple test sessions	64	21
Person to help child respond	‡	‡
Sign language interpreter	‡	‡
Other	15	3

‡ Reporting standards not met.
SOURCE: U.S. Department of Education, National Center for Special Education Research, Pre-Elementary Education Longitudinal Study (PEELS), "Parent Interview."

Table E-2. Unweighted number of children who had various test accommodations in the PEELS Wave 2 direct assessment, by race/ethnicity: School year 2004-05

	Black	Hispanic	White
Abacus	‡	‡	‡
Adaptive furniture	‡	3	7
Communication device	‡	‡	‡
Enlarged print	‡	‡	‡
Familiar person administered test	‡	‡	‡
Familiar person present	6	22	42
Multiple test sessions	9	14	56
Person to help child respond	‡	‡	‡5
Sign language interpreter	‡	‡	‡
Other	3	4	9

‡ Reporting standards not met.
NOTE: Some children who had accommodations are not included in this table because their race/ethnicity is not Black, Hispanic or White.
SOURCE: U.S. Department of Education, National Center for Special Education Research, Pre-Elementary Education Longitudinal Study (PEELS), "Parent Interview."

Table E-3. Unweighted number of children who had various test accommodations in the PEELS Wave 2 direct assessment, by primary disability: School year 2004-05

	AU	DD	ED	LD	MR	OI	OHI	SLI	LI
Abacus	‡	‡	‡	‡	‡	‡	‡	‡	‡
Adaptive furniture	‡	3	‡	‡	‡	3	‡	‡	‡
Communication device	‡	‡	‡	‡	‡	‡	‡	‡	‡
Enlarged print	‡	‡	‡	‡	‡	‡	‡	‡	‡
Familiar person present	14	21	‡	‡	5	‡	4	32	3
Multiple test sessions	9	28	‡	‡	3	‡	3	34	6
Person to help child respond	‡	‡	‡	‡	‡	‡	‡	‡	‡
Sign language interpreter	‡	‡	‡	‡	‡	‡	‡	‡	‡
Other	‡	4	‡	‡	3	‡	‡	3	4

‡ Reporting standards not met.
NOTE: AU = Autism; DD = Developmental delay; ED = Emotional disturbance; LD = Learning disability; MR = Mental retardation; OI = Orthopedic impairment; OHI = Other health impairment; SLI = Speech or language impairment; LI = Low incidence. Some children who had accommodations are not included in this table because they did not have a disability at the time the teacher questionnaire was administered; the teacher questionnaire was the source of the disability variable.
SOURCE: U.S. Department of Education, National Center for Special Education Research, Pre-Elementary Education Longitudinal Study (PEELS), "Parent Interview," "Early Childhood Teacher Questionnaire," and "Kindergarten Teacher Questionnaire."

Table E-4. Unweighted number of children who had various test accommodations in the PEELS Wave 2 direct assessment, by age cohort: School year 2004-05

	Cohort A (age 3)	Cohort B (age 4)	Cohort C (age 5)
Abacus	‡	‡	‡
Adaptive furniture	7	3	‡
Communication device	‡	‡	‡
Enlarged print	‡	‡	‡
Familiar person present	40	25	17
Multiple test sessions	26	36	23
Person to help child respond	3	3	‡
Sign language interpreter	‡	‡	‡
Other	4	7	7

‡ Reporting standards not met.
SOURCE: U.S. Department of Education, National Center for Special Education Research, Pre-Elementary Education Longitudinal Study (PEELS), "Parent Interview."

Appendix F: Analysis Variables Used Throughout Report

Variable	Source	Response codes
CHILD BACKGROUND AND FAMILY CHARACTERISTICS		
Age cohort	LEA sampling frame and parent interview	1=Cohort A 2=Cohort B 3=Cohort C
Child's sex	Parent interview	1=Male 2=Female
Race/ethnicity	Parent interview	1=Hispanic and of any race 2=Black or African American only, not Hispanic 3=White only and not Hispanic
Household income	Parent interview	1=$20,000 or less 2=$20,001-$40,000 3=More than $40,000
Disability category	Teacher questionnaire	1=Autism 2=Developmental Delay 3=Emotional Disturbance 4=Learning Disability 5=Mental Retardation 6=Orthopedic Impairment 7=Other Health Impairment 8=Speech or Language Impairment 9=Low incidence

Variable	Source	Response codes
SCHOOL /PROGRAM CONTEXT		
District wealth (Percent of district's children living in poverty)	QED sampling frame	1=High Wealth (0-12%) 2=Medium Wealth (13-34%) 3=Low Wealth (35-40%) 4=Very Low Wealth (>40%)
District size (Number of schools within the district)	QED sampling frame	1=Very large (391 or more) 2=Large (118-390) 3=Medium (42-117) 4=Small (41 or less)
Metropolitan status	QED sampling frame	1=Urban (large or mid-sized central city) 2=Suburban (Urban fringe of a large or mid-sized city, large or small town) 3=Rural (population of less than 2500)
Did child receive: adaptive physical education	Teacher questionnaire	1=yes 2=no
Assistive technology services/ devices	Teacher questionnaire	1=yes 2=no
Audiology	Teacher questionnaire	1=yes 2=no
Augmentative communication	Teacher questionnaire	1=yes 2=no

Variable	Source	Response codes
Behavior management program	Teacher questionnaire	1=yes 2=no
Health services	Teacher questionnaire	1=yes 2=no
Instruction in ASL	Teacher questionnaire	1=yes 2=no
Instruction in manual English or cued speech	Teacher questionnaire	1=yes 2=no
Instruction in Braille	Teacher questionnaire	1=yes 2=no
Learning strategies/study skills	Teacher questionnaire	1=yes 2=no
Mental health services	Teacher questionnaire	1=yes 2=no
Occupational therapy	Teacher questionnaire	1=yes 2=no
One-to-one paraeducator assistance	Teacher questionnaire	1=yes 2=no
Physical therapy	Teacher questionnaire	1=yes 2=no
Reader or interpreter	Teacher questionnaire	1=yes 2=no
Service coordination/case management	Teacher questionnaire	1=yes 2=no
Social work services	Teacher questionnaire	1=yes 2=no
Special transportation	Teacher questionnaire	1=yes 2=no
Specialized computer software	Teacher questionnaire	1=yes 2=no
Speech or language therapy	Teacher questionnaire	1=yes 2=no
Training, counseling, and other services to family	Teacher questionnaire	1=yes 2=no
Tutoring/remediation	Teacher questionnaire	1=yes 2=no
Vision services	Teacher questionnaire	1=yes 2=no
Other services	Teacher questionnaire	1=yes 2=no
Number of disability-related services	Teacher questionnaire	continuous, count of services received
Hours per week in regular education classroom	Teacher questionnaire	continuous
Hours per week in special education setting	Teacher questionnaire	continuous
Hours per week in therapy setting	Teacher questionnaire	continuous

Variable	Source	Response codes
Hours per week in nonspecial education setting outside regular class	Teacher questionnaire	continuous
Hours per week in home instruction	Teacher questionnaire	continuous

Variable	Source	Response codes
TRANSITIONS		
Location of enrollment the year before kindergarten	Teacher questionnaire	1=Exact same school and class as now 2=Same school but different kindergarten classroom 3=Not sure 4=Preschool class in same school 5=Some other program or at home
Transition status	Teacher questionnaire	1=No transition between Wave 1 and Wave 2 2=Transition between Wave 1 and Wave 2
OUTCOMES		
IEP/IFSP during previous year	Teacher questionnaire	1=yes 2=no
Declassified between Wave 1 and Wave 2 [must have been eligible at Wave 1]	Teacher questionnaire, missing data filled in using parent report	1=Has IEP/IFSP at both time points 2=IEP/IFSP at Wave 1 and declassified at Wave 2
Does child have an IEP?	Teacher questionnaire	1 = yes, has IEP/IFSP 2 = no IEP/IFSP
PKBS Problem Behaviors Scale	Teacher questionnaire	continuous variable
PKBS Social Skills Scale	Teacher questionnaire	continuous variable
PPVT	Child assessment	continuous variable
WJ Letter-Word Identification	Child assessment	continuous variable
WJ Applied Problems	Child assessment	continuous variable
WJ Quantitative Concepts	Child assessment	continuous variable
Child's academic skills compared to typical children of same grade level	Teacher questionnaire	1=Below or far below average 2= Average 3=Above or far above average